

ҒЫЛЫМИ-АНАЛИТИКАЛЫҚ ЖУРНАЛ
НАУЧНО-АНАЛИТИЧЕСКИЙ ЖУРНАЛ
SCIENTIFIC AND ANALYTICAL JOURNAL

ҚАЗАҚСТАННЫҢ ЖОҒАРЫ МЕКТЕБІ
ВЫСШАЯ ШКОЛА КАЗАХСТАНА
HIGHER EDUCATION IN KAZAKHSTAN

№3 (55) / 2026

ЖЫЛЫНА 4 РЕТ ШЫҒАРЫЛАДЫ
ВЫПУСКАЕТСЯ 4 РАЗА В ГОД PUBLISHED 4
TIMES A YEAR

2013 ЖЫЛДАН БАСТАП ШЫҒАДЫ
ИЗДАЕТСЯ С 2013 г.
FOUNDED SINCE 2013

Астана, 2026 жыл

Editor-in-chief

AIDOS MUKHATAYEV

Ministry of Science and Higher Education of the Republic of Kazakhstan,
Professor of the Astana IT University Candidate of Pedagogical Sciences, Associate
Professor (h-index – 10)

The owner:

Higher Education Development
National Center MSHE RK of
the Republic of Kazakhstan

Executive editor

YULIYA IDIYATOVA

EDITORIAL BOARD:

AMANTAY NURMAGAMBETOV - Adviser to the Director of the Higher
Education Development National Center, Ministry of Science and Higher Education
of the Republic of Kazakhstan, Doctor of Political Science, Professor

MARINA SKIBA - Higher Education Development National Center, Ministry of
Science and Higher Education of the Republic of Kazakhstan, Doctor of Political
Science, Professor (h-index – 2)

SERIK OMIRBAYEV - First Vice-Rector of QAIRU, Doctor of Economic Sciences,
Professor (h-index – 7)

NATALYA MUKAN - Professor at the Department of Pedagogy and Innovative
Education, National University "Lviv Polytechnic", Doctor of Pedagogical Sciences,
Professor (h-index – 2)

YURII ANDRASHKO – Associate Professor, Candidate of Technical Sciences,
Associate Professor at the Department of Systems Analysis and Optimization
Theory, Uzhhorod University, Ukraine (h-index – 22)

AINASH KUDYSHEVA - Vice-Rector for Academic and Educational-Methodical
Work, Kulyash Bayseyitova Kazakh National University of Arts, Candidate of
Pedagogical Sciences (h-index – 5)

SAPAR TOXANOV - Rector of QAIRU, PhD, (h-index - 6)

NIGINAKHON SHERMUKHAMEDOVA - Professor, Doctor of Philosophical
Sciences, National University of Uzbekistan named after Mirzo Ulugbek, Republic of
Uzbekistan (h-index – 1)

BAIBA RAMINA - Director of the Academic Information Centre, Latvia

MATTHEW A. ROSENSTEIN - Director of the Center for Global Education and
Learning, University of Illinois, Doctor of Philosophy (PhD), the United States of
America

Certificate No. 15650-Ж of
November 5, 2015 on
registration, rediscount of
periodical
printed publication and
informational agencies of the
Committee for
Communication,
Informatization and
Information of the Ministry for
Investments and
Development of the Republic
of Kazakhstan (primary
registration – Certificate No.
13306-Ж of January 25, 2013).

The journal is included in the
List of publications
recommended by the Science
and Higher Education Quality
Assurance Committee of the
Ministry of Science and Higher
Education of the Republic of
Kazakhstan for the publication
of the main results of scientific
activity (Order No. 374 of
March 26, 2024)

Signed for publication
30.09.2026.

CONTENTS

Bibigul Shagrayeva, Nailya Shertayeva, Mukhan Rakhmatulla, Bidos Tynabek EFFECTIVENESS OF THE STEM-BASED PROJECT METHOD IN TEACHING ORGANIC CHEMISTRY AT HIGHER EDUCATION INSTITUTIONS	6
Nurlan Mamiyev, Bakhytkhan Kuntugan, Arman Tukhmetov, Fakhrudin Batirov MOTIVATIONAL PROFILES AND PHYSICAL ACTIVITY OF UNIVERSITY STUDENTS: ASSOCIATIONS WITH LEARNING FORMAT AND SOCIAL SUPPORT	17
Tannur Bakytказы, Kuralay Nurgaliyeva, Eduarda Pinto Ferreira, Susana Nicola DEVELOPING STEM COMPETENCIES IN PRE-SERVICE PHYSICS TEACHERS: A LONGITUDINAL UNIVERSITY-BASED INTERVENTION STUDY IN KAZAKHSTAN	27
Aida Sagintayeva BEYOND REGULATION: INSTITUTIONAL CAPACITY AND DOCTORAL EDUCATION IN KAZAKHSTAN	39
Zhibek Tleshova, Nurgissa Kabarbek, Bexultan Turgunov THE ELECTRONIC SYLLABUS AS A PEDAGOGICAL OBJECT: AN INTEGRATIVE LITERATURE REVIEW	49
Zhanar Bekbayeva, Moldir Seidina, Saule Aldabergenova, Botagoz Bermukhambetova, Pawel Bogacz CRITICAL THINKING OF UNIVERSITY STUDENTS IN THE ERA OF ARTIFICIAL INTELLIGENCE	60
Damira Jantassova, Eleonora Prozovskaya EXPLORING THE IMPACT OF AI CHATBOTS ON ENGLISH-SPEAKING SKILLS IN ENGINEERING EDUCATION: A MIXED-METHODS STUDY	69
Askat Tleuov, Mukhametzhan Seitzhapparuly, Kairat Moldashev UNEVEN PARTNERSHIPS: PHD STUDENTS' SATISFACTION WITH LOCAL AND FOREIGN SUPERVISORY FEEDBACK AND INTERNATIONALIZATION OUTCOMES IN KAZAKHSTAN	80
Magzhan Amangeldi, Oleksandr Kuchanskyi FROM CITATION COUNTS TO COMPUTATIONAL SCIENCE: METHODS, DATA, AND APPLICATIONS IN SCIENTOMETRICS	93

Kamilya Arinova, Yerkin Nurmakhanov DIGITAL EDUCATIONAL ENVIRONMENT AS A RESOURCE FOR DEVELOPING DIGITAL COMPETENCIES OF FUTURE SPECIAL EDUCATORS: CHALLENGES AND PROSPECTS IN THE REPUBLIC OF KAZAKHSTAN	108
Lazura Kazykhankyzy, Meruyert Seitova, Ziyoda Khalmatova, H. Eray Çelik EXPLORING EFL TEACHERS' ATTITUDES: ADAPTATION AND VALIDATION OF THE PROFESSIONAL IDENTITY SCALE IN KAZAKHSTAN	122
Aziza Zhunusbekova, Samal Askarkyzy, Bakhyt Alipova AI ALGORITHMS FOR OPTIMIZING UNIVERSITY RESOURCE MANAGEMENT	137
Askat Tleuov, Temirlan Tursynbayev CURRICULUM INTERNATIONALISATION AND LOCAL ADAPTATION IN KAZAKHSTANI DUAL-DEGREE STEM PROGRAMMES	148

Bibigul Shagrayeva, Nailya Shertayeva*, Mukhan Rakhmatulla, Bidos Tynabek

O.Zhanibekov South Kazakhstan Pedagogical University, Shymkent, Kazakhstan

EFFECTIVENESS OF THE STEM-BASED PROJECT METHOD IN TEACHING ORGANIC CHEMISTRY AT HIGHER EDUCATION INSTITUTIONS

Abstract. The article examines the effectiveness of the project method based on STEM technology in teaching organic chemistry at universities. The relevance of the study stems from the complexity and importance of organic chemistry, as well as the need to increase students' interest in the subject and connect theoretical knowledge with practical application. To this end, the integration of project-based learning and STEM technologies creates conditions for the development of students' skills in research, creativity, and critical thinking. The purpose of the study is to show the effectiveness of the project approach based on STEM technologies in improving third-year students' mastery of organic chemistry at universities and to analyze the advantages and possible difficulties encountered in its application. The research used methods of theoretical analysis, pedagogical control, questionnaires, and comparative analysis of the results obtained. The findings show that the use of STEM technologies based on project-based learning in organic chemistry lessons increases students' learning motivation. In addition, this approach helps students develop practical skills and focus on future professional activities. Based on the results obtained, the article offers methodological recommendations for improving the learning process of organic chemistry.

Keywords: project-based learning, STEM technologies, organic chemistry, higher education institution, educational technologies, critical thinking, research skills.

Introduction

In his Address to the people of Kazakhstan "The economic benchmark of a just Kazakhstan", Head of State Kassym-Jomart Tokayev noted: "The formation of a new quality of the nation is of particular importance for our country. Every citizen, especially young people, should instill the best qualities. He came to the conclusion that everything together becomes a single public domain" (Shagrayeva, B. et al., 2025) The cited article presents research results showing that STEM-based learning improves conceptual understanding and practical skills, and develops key competencies in chemistry education. The consequences of the fourth technological revolution are visible in the modern scientific and educational space: high-speed information flows, high-tech innovations, and new developments are making changes in all spheres of modern society. In a rapidly developing world, the field of education is also undergoing significant changes. The main academic fields come to the fore, such as science, mathematics, technology, and engineering, which are united by a common word - STEM (science, technology, engineering, and mathematics) (Nogaibayeva G., 2026). STEM blocks affect almost all components of our daily lives. Today, STEM-related professions are among the fastest-growing and highest-paying professions, and have strong potential for employment growth. A specialist who has mastered any STEM profession has the appropriate competence, so the abbreviation STEM is firmly "entrenched" in the educational space, and the term "STEM education" has emerged (Karatayeva M. S. et al., 2023).

In the Republic of Kazakhstan, special attention is paid to the introduction of STEM technologies in universities as a part of the transition to a learning model based on updated educational content. STEM technologies allow students to apply theoretical knowledge in practice, which is particularly important in organic chemistry and related applied disciplines. Students immerse themselves in the creative process with the help of STEM technologies. there they learn new information based on their accumulated knowledge. In addition, students can rely on teacher support when solving practical or scientific problems. As a result, participation in such activities helps students acquire new knowledge (Konyushenko S.M. et al., 2017).

In recent years, the method of project-based learning has been widely used in the field of education. Project-based learning is based on students' independent search, problem-solving, and activities aimed at achieving specific results. The project method related to STEM technologies, in addition to deepening students'

subject knowledge, contributes to the formation of their thinking and creative abilities as engineers. However, the effectiveness of the project-based approach based on STEM technologies in teaching organic chemistry in universities has been systematized at a fairly high level and has not been scientifically analyzed. Innovation in learning is crucial to create a new and enjoyable learning experience for students. Innovation in teaching has been proven to improve learning outcomes by providing innovative learning resources. Research also indicates that project-based learning promotes active learning and competency development. It allows students to strengthen their analytical and problem-solving skills. (Styers M L. et al., 2018).

Despite the widespread adoption of project-based learning (PBL) and STEM education in contemporary higher education, their integrated application in teaching organic chemistry remains under-researched. In particular, limited attention has been paid to instructional formats that combine project-based quest activities, molecular modeling, Arduino-based tasks, digital tools, and collaborative problem-solving within a unified learning model. This gap is particularly evident in university-level chemistry education, where abstract concepts of molecular structure must be bridged with practical and interdisciplinary activities. Therefore, this study aims to develop and pilot an integrated STEM-based project learning model for organic chemistry.

Literature review

The abbreviation STEM (Science, Technology, Engineering, Mathematics) was proposed in the mid-1990s by the American bacteriologist Rita Colwell in 2001 was introduced by the US National Science Foundation into wide rotation to denote an educational system in which individual subjects (physics, chemistry, biology, mathematics, etc.) were not taught. From various sciences (primarily natural sciences and engineering), ways to solve the problems of attracting knowledge. In the vast majority of cases, the "points of integration" of this knowledge are projects. In this regard, STEM education acts as the development of a project approach. But in contrast, it necessarily includes the study of at least several scientific fields (for example, not only physics, but also biology, programming) at the basic level of subjects – the set of these subjects may vary depending on the student's specialization. As a rule, working on a project is combined with an internship at a specific technology company, so real business problems are often addressed during the training process. Although STEM is an abbreviation, the word "stem" also has an independent meaning in English, as in the phrase "stem cell." This type of education encourages students to pursue lifelong learning. Recalling the familiar analogy of giving a person a fish or a fishing rod, STEM education provides neither one directly; instead, it develops the ability to make the "fishing rod" independently. (Zhumadullayeva A. A. et al., 2025).

In the first decades of the 21st century, STEM education became widespread in leading Western countries. Its principles, approaches, and methods can be found on open online educational platforms: Coursera, edX, and Udemy, where universities around the world publish courses of varying duration and complexity. Since 2016, the European Journal of STEM Education has been published in the Netherlands. The main part of his materials is devoted to the practical implementation of STEM education in European universities and schools. Since 2018, China has been implementing the "Chinese Innovative Action Plan for STEM Education until 2029", which provides for fundamental changes in the field of professional training. The officially announced goal of transitioning to STEM education requires consistent changes in educational standards and curricula, starting at the school level.

The study by Sahito and Wassan (2024) explores challenges, issues, and solutions in the natural sciences, with a focus on STEM education. The researchers analyze key aspects, including the objectives and complexities of STEM education, curriculum, inquiry, learning environments, technologies, learning models and integration, blended learning approaches, students' problem-solving skills, implementation and evaluation, research, current trends, and interdisciplinary collaboration. The renewal of education should be based on the methodology of designing educational systems, on values and goals that support the nature of various educational institutions and their attempts at self-development. Design as a promising direction for the development of education is studied by V.I.Slobodchikov from the point of view of his innovative activity in education. The project approach is a way of learning in which students join groups and work on some financially secure, income-generating project. In this case, the knowledge gained by students does not depend on the structure of scientific disciplines, but on the problems that arise during the development of a specific project. The project approach has been used in education for a long time. In particular, project-based teaching

methods were actively introduced in the Soviet school of the 1920s. In Europe, at the beginning of the twentieth century, the technology of project-based learning in the field of education was formed. In Kazakhstan, the technology of project-based learning in the field of education has been applied since the introduction of the credit training system (Langie, G. et al., 2018). Project-based learning technologies are also widely used in the education system of the CIS countries. The first to develop the theoretical basis of project-based learning technology were Russian scientists V. P. Bespalko, V. V. Davydov, V. K. Dyachenko, L. V. Zankov, P. Ya. Galperin, N. V. Kuzmina. As a result of the theoretical research of these scientists, E. N. Ilyina, S. N. Lysenkova, V. F. Shatalov, and T. S. Methodologists conducted research that is widely used in their practical experiments (Angelova, O., et al., 2020).

The basic principles of project activity are significant and are defined in STEM education as the nature of research: research ability; independence; practical orientation (Blesman A.I., et al., 2021). The concept of STEM education is broader than the basic abbreviation, as it allows students to holistically perceive the world around them, eliminating traditionally established barriers between the main academic disciplines included in STEM education (Sologub N.S., et al., 2020).

The advantages and disadvantages of STEM education are related to the fact that it is a complex phenomenon based on a practice-oriented approach that is actively used to prepare students for future life based on the requirements of the labor market. The implementation of STEM education is characterized by both advantages and difficulties due to the high costs of retraining or advanced training of teachers and others (Medina V., et al., 2021). When implementing STEM education, humanities subjects may be underestimated. Therefore, the successful implementation of STEM education, including through project activities, requires a balanced approach, taking into account both its strengths and possible limitations (Yang H., et al., 2024). Project activities in STEM education are based on a variety of methods and approaches that help students acquire and apply knowledge in a practical environment. Let's define the main ones: design method; engineering design; query-based learning (IBL); five stages of the design model; active learning methods; collaborative learning; interactive technologies; game methods; specific projects. All of them have their own characteristics and are combined with the fact that they aim to allow students to participate more actively in the learning process, which allows them to solve specific problems and apply their knowledge in practice. A combination of various methods is often used to create a universal learning space (Saleh S., 2020).

Furthermore, the STEM approach boosts students' engagement, cognitive skills, and creativity. STEM education enables learners to manage the vast flow of information, apply their knowledge practically, take initiative, and work confidently with equipment and technology (Ryu et al., 2019). By studying the basic principles of the natural sciences, students gain an understanding of the links between different phenomena and uncover new scientific facts (Chiu, 2023). Additionally, it is important to highlight the vital role of STEM education in fostering and maintaining students' interest in STEM careers, as it helps shape their career goals within the STEM field (Ribeirinha et al., 2024).

Previous research indicates that STEM education and project-based learning (PBL) can enhance student engagement, foster collaboration, improve problem-solving skills, and deepen conceptual understanding. However, most studies examine these approaches in isolation or focus primarily on K-12 science education. Significantly less attention has been paid to the integration of molecular modeling, digital technologies, Arduino-based hands-on tasks, and quest-based project methods in university-level organic chemistry courses. Furthermore, existing literature lacks sufficient data on the feasibility of implementing such an integrated model for chemistry majors.

Thus, the research gap lies in the scarcity of empirically documented examples of a comprehensive STEM-based project approach specifically designed for higher education organic chemistry. This study aims to address this gap by developing and piloting a structured learning model that integrates theoretical chemistry knowledge, digital modeling, practical technological tasks, and collaborative project-based activities.

Research methods and materials

The study was conducted in the 2025-2026 academic year and consisted of three stages: preparatory, practical, and final. At the preparatory stage, the analysis of scientific and pedagogical literature related on STEM technologies and the method of project-based learning in teaching organic chemistry was carried out, the purpose and objectives of the research were specified, and survey tools were developed.

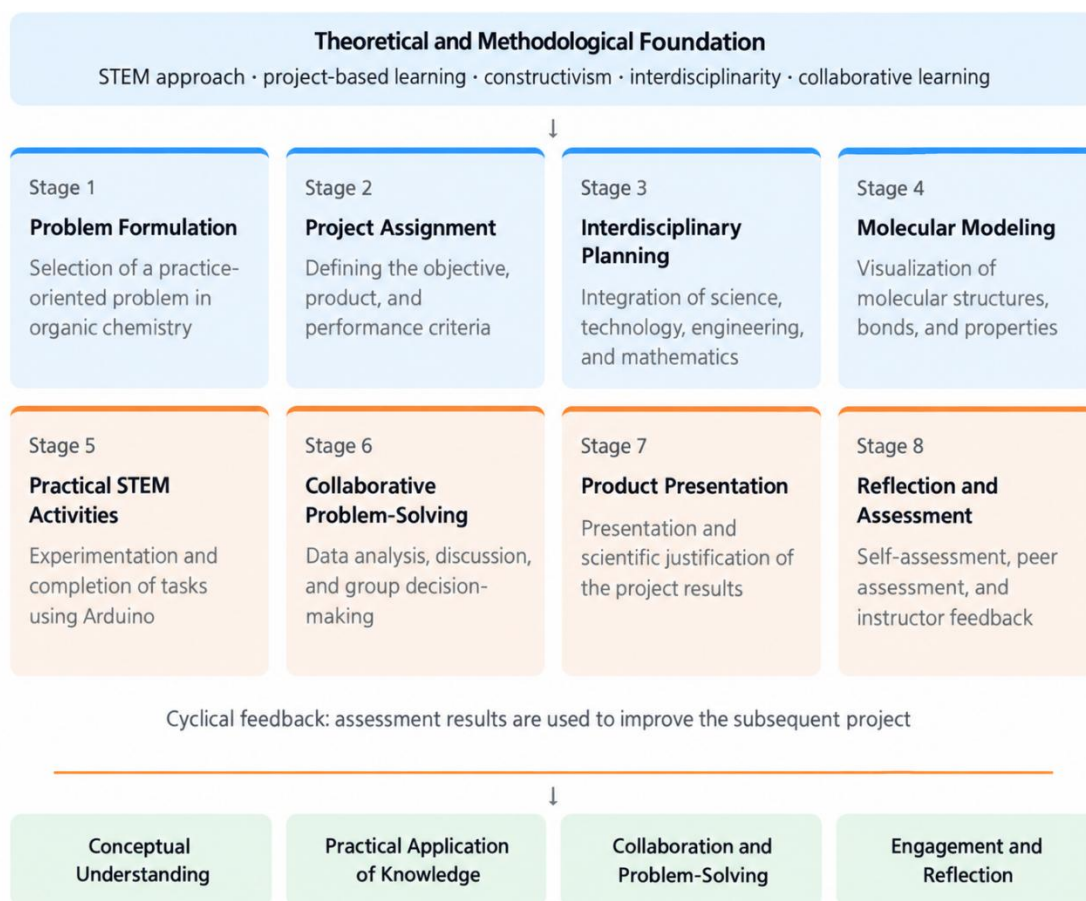
The study involved 50 students who were divided into two equal groups: a control group (CG, $n = 25$) and an experimental group (EG, $n = 25$). In the experimental group, instruction was conducted using the developed methodology, whereas the control group followed a traditional approach. To evaluate the effectiveness of the methodology, both groups underwent pre-test and post-test assessments.

The study was conducted in accordance with generally accepted ethical principles for educational research. Participants were informed about the purpose, procedure, and conditions of the study. Informed consent was obtained from the participants prior to the commencement of the research. Confidentiality was ensured through data anonymization and the use of individual codes instead of personal identifiers. The results obtained were used exclusively in an aggregated form and for scientific purposes.

In the experimental group, instruction was based on the developed integrated STEM project-based learning model for organic chemistry. This model integrates the content of natural science and technical disciplines, project-based activities, the solution of practice-oriented tasks, and the continuous assessment of learning outcomes. The structure and key components of the developed model are presented in Figure 1.

Figure 1

Integrated STEM Project-Based Model for Teaching Organic Chemistry



The implementation of the model comprised four stages: motivational-goal-oriented, content-activity-based, project-practical, and evaluative-reflective. At each stage, students applied knowledge from chemistry, biology, physics, mathematics, and technology to solve complex academic and practical problems.

Statistical data analysis was performed on a sample of 50 participants, divided into a control group ($n = 25$) and an experimental group ($n = 25$). Descriptive statistics included the mean (M), standard deviation (SD), median (Me), interquartile range (IQR), and 95% confidence intervals. The normality of the distribution was assessed using the Shapiro-Wilk test, and the homogeneity of variances was verified using Levene's test. Within-group changes were analyzed using the paired-samples t -test and the Wilcoxon signed-rank test, while between-group comparisons were conducted using Welch's t -test and the Mann-Whitney U test. The intervention effect was further evaluated using ANCOVA, controlling for pre-test scores. Effect sizes were

determined using Cohen's d , Hedges' g , and partial η^2 . Differences were considered statistically significant at $p < 0.05$.

The results of the study

The study experimentally evaluated the effectiveness of project-based learning for 3rd-year university students based on STEM technologies in the study of organic chemistry. The results of the study were determined by comparing the formats of project and quest lessons related to traditional lecture lessons and STEM technologies.

The practical lessons showed that the STEM-based project approach significantly increased students' active participation in the learning process. During the project-quest sessions, students were involved in group work and studied organic chemistry topics in depth through building molecular models, analyzing assignments, and solving practical problems. During the classes, the students actively exchanged opinions, asked questions, and made decisions.

The survey results showed that teaching organic chemistry to the vast majority of students using a project-based approach based on STEM technologies is more interesting and understandable than a traditional lecture lesson. The students stated that this approach allowed them to connect theoretical knowledge with real-life. Thus, the project assignments helped to increase the students' learning motivation.

The results of a comparative analysis of academic achievements showed that students' level of assimilation of the topic increased after STEM-based project classes. Compared with traditional lecture classes, lessons that used a project approach produced higher scores for assignment quality, the accuracy of answers to theoretical questions, and practical problem-solving. This proves that the project approach based on STEM technologies is an effective method in mastering organic chemistry.

The study also showed positive changes in the development of students' research and creative abilities. Project assignments encouraged students to search for information, analyze data, and defend their results. This contributed to the development of their critical thinking skills and increased their level of preparation for future professional activities.

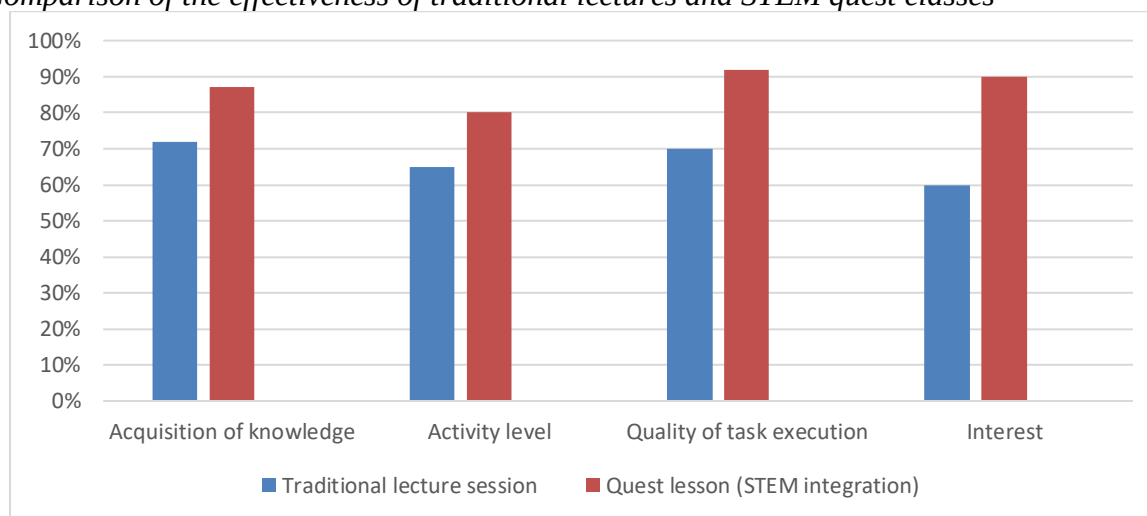
In general, the results of the study showed that a project-based approach based on STEM technologies in teaching organic chemistry is effective in improving the quality of students' education, increasing motivation to learn, and developing research skills. The findings indicate the pedagogical value of systematically implementing STEM technologies in the process of teaching organic chemistry in universities.

Table 1

Comparative analysis of the results of classes

Criteria	Traditional lecture session	Quest lesson (STEM integration)
Number of students involved	25 students	25 students
Lesson format	Lecture, theoretical presentation	Quest, teamwork, interactive stations
Acquisition of knowledge	Average score: 72/100	Average score: 87/100
Activity level	Low – 35%	High – 80%
Answers to questions	Personal response (passive listening)	Group response, discussion
Interest	Average	High
Impressions of the lesson	Feedback "I got a lot of information"	Feedback "was interesting", "was unforgettable"
Quality of task execution	60-80%	85-95%
Reflection of 21st-century skills	Limited (listening, recording)	Fully covered (collaborative work, research, logic, decision-making)

The level of student participation in the lectures was also determined and included in the scheme (Figure 2).

Figure 2*Comparison of the effectiveness of traditional lectures and STEM quest classes*

The diagram below compares the results of traditional lectures and STEM quest lessons based on four indicators: knowledge acquisition, activity level, task quality, and interest. First, the average knowledge-acquisition score was 72% during traditional lectures and 87% during STEM quest lessons. This suggests that the use of STEM elements contributed to better knowledge acquisition. Second, the largest difference was observed in activity level: 65% in traditional lectures compared with 80% in STEM quest lessons. This proves that tasks in the quest format effectively contribute to independent work and increase students' motivation. Third, task quality was also substantially higher in the STEM quest lesson: 70% in the traditional lesson compared with 92% in the quest lesson. This indicates that STEM methods increase students' responsibility, creativity, and accuracy. Fourth, the interest rate reached 60% in the traditional lecture and 90% in the STEM quest. This result suggests that engaging lesson organization plays an important role in students' interest.

The quest lesson, conducted through the integration of STEM technology, yielded more effective results in all respects compared to a traditional lecture. Students' activity and interest increased particularly noticeably. This shows that they are more deeply involved in the learning process and have better assimilated the acquired knowledge.

To evaluate the effectiveness of the developed model, the results of the pre-test and post-test for both the control and experimental groups were compared. In the first stage, descriptive statistics were calculated (Table 2), followed by intra-group and inter-group comparisons, as well as an analysis of covariance (ANCOVA) controlling for the participants' baseline levels.

Table 2*Pre-test, post-test, and gain score descriptive statistics for control and experimental groups*

Group	Measure	n	M	SD	Me	IQR	Min–ax
Control group	Pretest	25	67,12	2,09	67,00	4,00	64–71
Control group	Posttest	25	71,12	1,94	71,00	2,00	67–75
Control group	Gain score	25	4,00	1,41	4,00	2,00	0–6
Experimental group	Pretest	25	68,28	2,54	68,00	3,00	64–73
Experimental group	Posttest	25	86,96	3,72	86,00	5,00	79–93
Experimental group	Gain score	25	18,68	3,22	18,00	3,00	14–28

Note. n = sample size; M = mean; SD = standard deviation; Me = median; IQR = interquartile range; Min–Max = minimum and maximum values.

Before the experiment, no statistically significant differences were found between the control and

experimental groups: the mean pre-test scores were 67.12 and 68.28 points, respectively ($p = 0.084$). Therefore, the groups can be considered comparable in terms of their baseline levels of the measured outcome.

Table 3

Results of the Shapiro–Wilk Test for Normality of Score Distributions in the Control and Experimental Groups

Group	Measure	W Shapiro–Wilk W	p
Control group	Pretest	0,948	= 0,222
Control group	Posttest	0,949	= 0,237
Control group	Gain score	0,834	< 0,001
Experimental group	Pretest	0,943	= 0,177
Experimental group	Posttest	0,944	= 0,187
Experimental group	Gain score	0,851	= 0,002

As shown in Table 3, unequal variances were identified for the gain scores ($p = 0.048$). Therefore, Welch's t-test was used to compare the control and experimental groups.

Table 4

Results of Levene's Test for Homogeneity of Variances in Pretest, Posttest, and Gain Scores

Measure	F Levene's	p
Pretest	0,731	= 0,397
Posttest	8,159	= 0,006
Gain score	4,126	= 0,048

Levene's test results (Table 4) showed that for the **gain** metric, the differences in variances between the control and experimental groups reached a statistically significant level ($p = 0.048$). Since the p-value is less than 0.05, the assumption of homogeneity of variances for this metric is not met. Therefore, when comparing the **gain** between groups, it is advisable to use Welch's t-test, which does not assume equal variances. For the **pre-test** and **post-test** metrics, the interpretation is analogous: with $p > 0.05$, variances are considered homogeneous, and with $p < 0.05$, they are considered heterogeneous.

Table-5

Results of the Within-Group Comparison of Pretest and Posttest Scores in the Control and Experimental Groups

Group	ΔM	95% CI	t	df	p	d^z	W	p(Wilxon)
Control group	4,00	3,42; 4,58	14,14	24	< 0,001	2,83	0,0	< 0,001
Experimental group	18,68	17,35; 20,01	28,97	24	< 0,001	5,79	0,0	< 0,001

In both groups, a statistically significant improvement in scores was observed after training ($p < 0.001$). In the control group, the mean increase was 4.00 points (95% CI: 3.42–4.58; $t(24) = 14.14$; $d^z = 2.83$), whereas in the experimental group, it was 18.68 points (95% CI: 17.35–20.01; $t(24) = 28.97$; $d^z = 5.79$). The Wilcoxon signed-rank test also confirmed the significance of the changes in both groups ($W = 0$; $p < 0.001$). Thus, the most pronounced improvement in scores was observed in the experimental group.

Table 6

Findings from the intergroup comparison of pre-test, post-test, and gain scores for the control and experimental groups

Measure	ΔM	95% CI	Welch's t-test	df	p	Hedges' g	U	p(Mann-Whitney test)
Pretest	1,16	-0,16; 2,48	1,76	46,26	= 0,084	0,49	390,5	= 0,129
Posttest	15,84	14,14; 17,54	18,85	36,17	< 0,001	5,25	625,0	< 0,001
Gain score	14,68	13,25; 16,11	20,85	32,91	< 0,001	5,80	625,0	< 0,001

Note. The difference was calculated as the experimental group minus the control group

Before the intervention, no statistically significant differences were found between the groups: $\Delta M = 1.16$ points, 95% CI $[-0.16, 2.48]$, $t(46.26) = 1.76$, $p = 0.084$; the Mann-Whitney U test was also nonsignificant ($p = 0.129$). After the intervention, the experimental group outperformed the control group by 15.84 points, 95% CI $[14.14, 17.54]$, $t(36.17) = 18.85$, $p < 0.001$, Hedges' $g = 5.25$. The between-group difference in gain scores was 14.68 points, 95% CI $[13.25, 16.11]$, $t(32.91) = 20.85$, $p < 0.001$, Hedges' $g = 5.80$. The nonparametric analysis confirmed significant between-group differences in both post-test and gain scores (both $p < 0.001$).

ANCOVA results for posttest scores, controlling for baseline outcome (mean pretest = 67.70), revealed an adjusted posttest mean of 71.55 for the control group and 86.53 for the experimental group. The adjusted advantage of the experimental group was 14.98 points, 95% CI $[13.54; 16.41]$, $F(1,47) = 441.66$, $p < 0.001$, partial $\eta^2 = 0.904$. The "group $\times$ pre-test" interaction was nonsignificant, $F(1,46) = 0.04$, $p = 0.834$, indicating that the assumption of homogeneity of regression slopes was met.

The experimental and control groups were statistically comparable in their pre-test results. Both groups improved, but the experimental group showed a significantly greater gain of 14.68 points. This conclusion is robust, as it is supported by Welch's t-test, the Mann-Whitney U test, and ANCOVA adjusted for baseline. The obtained data indicate the strong effectiveness of the experimental training in this sample. However, the unusually large effect sizes should be interpreted in light of the study design, the method of participant assignment to groups, and the characteristics of the measurement instrument; statistical significance alone does not prove causality in the absence of randomization.

Analysis of the research results

The study results indicated that the implementation of the integrated STEM project-based learning model, presented in Figure 1, contributed to a more pronounced improvement in the academic achievements of students in the experimental group. The average gain in scores in the experimental group was 18.68 points, while in the control group it was 4.00 points. The intergroup difference in the magnitude of the gain was 14.68 points and was statistically significant ($p < 0.001$).

The identified advantage of the experimental group may be linked to the comprehensive structure of the integrated STEM project model (Figure 1). The combination of interdisciplinary content, project-based learning, practice-oriented tasks, and reflective assessment ensured the active engagement of students in the learning process. Unlike traditional teaching, this model involved the application of organic chemistry knowledge in solving practical problems, developing projects, and interpreting the obtained results. This could have contributed to a deeper understanding of the learning material and a higher gain in the experimental group's outcomes.

Conclusion

This pilot study examined the practical applicability of an integrated STEM-based project model for

teaching organic chemistry at the university level. The model combined project- and quest-based activities, molecular modeling, digital tools, Arduino-based tasks, and collaborative problem-solving. The findings indicate that this approach may enhance students' engagement and interest, foster collaboration, and facilitate the practical application of theoretical knowledge.

At the same time, the findings should not be regarded as conclusive evidence that STEM-based project learning is more effective than traditional instruction. The small sample size, the single-university setting, the insufficiently detailed description of the comparison conditions, and the predominant use of descriptive measures limit the generalizability of the findings. Therefore, this study should be regarded as exploratory research that provides a foundation for subsequent methodologically rigorous investigations.

The main contribution of this study lies in the development and initial testing of an integrated instructional model adapted to a university-level organic chemistry course. Future research should evaluate this model using larger samples, reliable assessment instruments, a controlled experimental design, and appropriate statistical methods.

Funding Information

This research was conducted as part of the project BR28713505 titled «Development of Conceptual Foundations for Advancing Rural and Small-Scale Schools in the Republic of Kazakhstan» and of the grant project AP26102613 titled «Development of methodological aspects of the application of virtual and augmented reality technologies in teaching chemistry and biology in the context of implementing an integrative approach in the educational process» funded under the targeted program for scientific and scientific-technical development for 2025–2027 by the Ministry of Science and Higher Education of the Republic of Kazakhstan.

Conflict of Interest Statement

The authors declare no potential conflicts of interest regarding the research, authorship, or publication of this article.

Declaration of the use of AI

The authors declare that no artificial intelligence (AI) tools were used in the preparation of this manuscript.

Author Contributions

Bibigul Shagrayeva: Conceptualization, Methodology, Writing – Original Draft Preparation, Supervision, Editing. Mukhan Rakhmatulla: Theoretical Analysis, Literature Review, Writing – Reviewing and Editing. Bidos Tynabek: Project Administration, Resources, Literature Review, Writing – Reviewing and Editing. Nailya Shertayeva: Writing – Review and Editing.

References

- Angelova, O., et al. (2020). Typology of mentor's professional activity styles. *Proceedings of the Southwest State University*, 10(6), 220–228.
- Blesman, A. I., Poleshchenko, K. N., Semenyuk, N. A., & Teplouhov, A. A. (2021). *Osnovy proektnoi deyatel'nosti: Metodicheskie ukazaniya* [Basics of project activity: Methodical instructions]. Omsk State Technical University.
- Chiu, T. K. F., & Li, Y. (2023). How can emerging technologies impact STEM education? *Journal of STEM Education Research*, 6, 375–384. <https://doi.org/10.1007/s41979-023-00113-w>
- Jumadullaeva, A. A., Sultanbek, M. J. Q. S., & Nurahmetova, Z. A. (2025). Jobalau ädısı arqyly oquşylardyñ äleumettenuın qamtamasyz etudiñ diagnostikasy [Diagnostics of ensuring students' socialization through the project-based method]. *Habarshysy of Abai Kazakh National Pedagogical University. Pedagogical Sciences Series*, 87(3), 359–375.
- Kazbekova, G. N., & Ismagulova, Zh. S. (2022). Innovatively STEM-bilim beru tasilin qalyptastyru [Formation of innovative STEM education]. *Iasau University Bulletin*, 3(125), 200–210. <https://doi.org/10.47526/2022-3/2664-0686.17>

- Kusainova, S. M., Tantybaeva, B. S., Abylaikhan, A., Shaiikhova, B. K., & Erejepekova, A. A. (2023). Ispol'zovanie innovatsionnykh metodov STEM na urokakh khimii s realizatsiei mezhpredmetnykh svyazei [The use of innovative STEM methods in chemistry lessons with interdisciplinary integration]. *Bulletin of L. N. Gumilev Eurasian National University. Pedagogy. Psychology. Sociology Series*, 143(2), 212–220.
- Langie, G., & Pinxten, M. (2018). The transition to STEM higher education: Policy recommendations—Conclusions of the Ready STEM Go project. *International Journal of Engineering Pedagogy*, 8(10), 10–13.
- Medina, V., & Todd, R. (2021). Implementing STEM-related activities for school-aged users. *IASL Annual Conference Proceedings*, 1–8. <https://doi.org/10.29173/iasl8285>
- Nogaibaeva, G. (2026). Razvitie STEM-obrazovaniya v mire i Kazakhstane [Development of STEM education in the world and Kazakhstan]. <https://iac.kz/ru/publishing/razvitie-stem-obrazovaniya-v-mire-i-kazakhstane>
- Ontagarova, D. R., & Zholymbayev, O. M. (2023). STEM oqytudyñ ädisteri men formalary [Methods and forms of STEM education]. *Kazakh University*.
- Qarataeva, M. S., Berkimbaev, K. M., & Jaidakbaeva, L. K. (2023). STEM bilim beruge bolaşaq informatika mūğalımdırın daıarlaudyñ mazmūny [Content of training future computer science teachers for STEM education]. *Bulletin of L. N. Gumilev Eurasian National University. Pedagogy. Psychology. Sociology Series*, 145(4), 110–120.
- Ribeirinha, T., Baptista, M., & Correia, M. (2024). The impact of STEM activities on the interest and aspirations in STEM careers of 12th-grade Portuguese students in science and technology curriculum. *European Journal of STEM Education*, 9(1), 21. <https://doi.org/10.20897/ejsteme/15830>
- Ryu, M., Mentzer, N., & Knobloch, N. (2019). Preservice teachers' experiences of STEM integration: Challenges and implications for integrated STEM teacher preparation. *International Journal of Technology and Design Education*, 29(3), 493–512. <https://doi.org/10.1007/s10798-018-9440-9>
- Sahito, Z., & Wassan, S. H. (2024). Literature review on STEM education and its awareness among teachers: An exploration of issues and problems with their solutions. *SAGE Open*, 14(1). <https://doi.org/10.1177/21582440241236242>
- Saleh, S., Muhammad, A., & Abdullah, S. (2020). STEM project-based approach in enhancing conceptual understanding and inventive thinking skills among secondary school students. *Journal of Nusantara Studies*, 5, 234–254. <https://doi.org/10.24200/jonus.vol5iss1pp234-254>
- Shagraueva, B., Shertayeva, N., Zhorabekova, A., Baymakhanova, G., & Toktamys, N. (2025). Enhancing the quality of education through the integration of STEM technology elements in chemistry lessons. *European Journal of STEM Education*, 10(1), 15. <https://doi.org/10.20897/ejsteme/17176>
- Sologub, N. S., & Arshanskii, E. Ya. (2020). STEAM-obrazovanie: Sushchnost' i analiz idei v istoricheskoi retrospektive [STEAM education: Essence and analysis of ideas in historical retrospect]. *Vesci BDPU*, 2, 15–18.
- Styers, M. L., Van Zandt, P. A., & Hayden, K. L. (2018). Active learning in flipped life science courses promotes development of critical thinking skills. *CBE—Life Sciences Education*, 17(3), Article ar39. <https://doi.org/10.1187/cbe.16-11-0332>
- Yang, H., Zhang, Q., & Shen, M. (2024). The practice and research of junior high school information technology project-based learning based on STEM education concept. *International Journal of Technology in Education and Science*, 8(1), 63–74. <https://doi.org/10.46328/ijtes.537>
- Yesnazar, A., Zhorabekova, A., Kalzhanova, A., Zhuzimkul, B., & Almukhanbet, S. (2024). Methodological system for the formation of meta-subject skills of primary school students in the context of STEM education. *Frontiers in Education*, 9, Article 1340361. <https://doi.org/10.3389/feduc.2024.1340361>

Information about authors

Bibigul Shagrayeva - acting professor, candidate of chemical sciences, South Kazakhstan Pedagogical University named after O.Zhanibekov, Shymkent, Kazakhstan, e-mail: Bibi-0305@mail.ru, ORCID 0000-0002-9606-8709

Nailya Shertayeva - acting professor, candidate of chemical sciences, South Kazakhstan Pedagogical University named after O.Zhanibekov, Shymkent, Kazakhstan, e-mail: Nailyaximik@mail.ru, ORCID 0000-0001-6446-4953
(corresponding author)

Mukhan Rakhmatulla – master's student in Chemical sciences, South Kazakhstan Pedagogical University named after O. Zhanibekov, Shymkent, Kazakhstan, e-mail: mmukhan03@gmail.com, ORCID 0009-0006-5097-2707

Bidos Tynabek - master's student in Chemical sciences, South Kazakhstan Pedagogical University named after O.Zhanibekov, Shymkent, Kazakhstan, e-mail: btynabek@bk.ru, ORCID 0000-0002-2252-3270

Received: 31.01.2026

Revised: 26.07.2025

Revised: 31.08.2026

Accepted: 15.09.2026

Published: 30.09.2026

Type of article: Research Article

Nurlan Mamiyev*, Bakhytkhan Kuntugan, Arman Tukhmetov, Fakhridin Batirov

*Department of Physical Education, S. Seifullin Kazakh Agrotechnical Research University,
Astana, Kazakhstan.*

MOTIVATIONAL PROFILES AND PHYSICAL ACTIVITY OF UNIVERSITY STUDENTS: ASSOCIATIONS WITH LEARNING FORMAT AND SOCIAL SUPPORT

Abstract. The relevance of this study is determined by the need to understand how motivational regulation and university learning conditions are associated with students' regular physical activity. The aim was to identify motivational profiles and examine their associations with physical activity, learning format, and social support. The quantitative observational study included 250 first- and second-year students (185 men and 65 women; 17–20 years) enrolled in the 6B07102 Electric Power Engineering program at S. Seifullin Kazakh Agrotechnical Research University. A 12-item author-developed questionnaire informed by self-determination theory assessed autonomous regulation, controlled regulation, social support, and subjective appraisal of physical activity. The archived analysis distinguished three descriptive motivational profiles. Weekly physical activity differed substantially across the profiles, $F(2, 247) = 149.17$, $p < .001$, $\eta^2 = .547$: autonomous-profile students reported 4.60 ± 1.20 active days per week, controlled-profile students 3.20 ± 1.00 , and low-motivation students 1.40 ± 0.70 . Social support was positively correlated with physical activity ($r = .35-.48$, $p < .001$). Descriptive differences were also observed across learning formats. The findings identify associations rather than causal effects and support motivation-sensitive and socially supportive approaches to physical education in higher education.

Key words: university students, physical activity, motivational profiles, self-determination theory, social support, learning format.

Introduction

In contemporary research at the intersection of education and health, students' physical activity is increasingly conceptualized as a key determinant of human capital formation, academic achievement, and a sustainable lifestyle in the long term. Amid the transformation of the university educational environment driven by digitalization, the expansion of hybrid learning formats, and shifts in patterns of social interaction the challenge of maintaining and developing systematic physical activity among learners has become distinctly interdisciplinary, extending beyond the boundaries of traditional physical education.

In the scholarly literature, students' physical activity is progressively examined through the lens of motivational frameworks, most notably self-determination theory, which posits that the stability of behavioral practices is associated with the predominance of autonomous rather than controlled motivation (Ryan & Deci, 2000). Empirical evidence and systematic reviews further indicate that intrinsic motivation and the satisfaction of basic psychological needs autonomy, competence, and relatedness serve as robust predictors of regular physical activity in student populations (Teixeira et al., 2012; Ntoumanis et al., 2021). At the same time, the educational environment is emphasized as a contextual factor that can either support or constrain students' motivational resources.

In recent years, increasing attention has been devoted to the transformation of the university educational environment as an integrated set of spatial, organizational, digital, and social-communicative conditions shaping students' behaviors and attitudes. Studies addressing changes in lifestyle and interaction patterns during periods of intensified digitalization and pandemic-related restrictions point to an ambivalent effect: on the one hand, flexibility and personalization of the educational process are strengthened; on the other, the risks of insufficient physical activity and increased sedentary behavior rise, potentially accompanied by weakened everyday social ties (Guthold et al., 2020; Hall et al., 2021). Accordingly, the purposeful design of an educational environment oriented not only toward academic outcomes but also toward the maintenance of students' physical activity and well-being becomes increasingly critical.

For the higher education system of the Republic of Kazakhstan, the relevance of this issue is further

amplified by institutional transformations, the implementation of digital educational platforms, and a reconsideration of universities' roles in fostering a healthy and socially active generation. Despite the existence of individual studies on the physical activity of university youth, the field still lacks a sufficiently systematized analysis of the relationships among students' motivational characteristics, features of the transformed educational environment, and social factors, including team dynamics and forms of support. This gap necessitates a comprehensive study integrating motivational, environmental, and social parameters.

The aim of this study was to identify students' motivational profiles and examine their associations with physical activity, learning format, and social support in the university context.

To achieve this aim, the study addressed the following objectives:

- to critically compare theoretical and empirical approaches to motivation for university students' physical activity;
- to describe physical-activity indicators across face-to-face, hybrid, and distance learning formats;
- to characterize students' motivational profiles and compare their weekly physical activity;
- to examine the associations of social support and team-based participation with physical activity;
- to interpret the findings as observational associations and identify implications for university physical-education practice.

The theoretical significance of the study lies in integrating self-determination theory with learning-context and social-support indicators within one university sample. Its practical significance lies in identifying motivational and social patterns that may inform the design of physical-education and student-support practices without treating observational associations as causal effects.

Literature review

Self-determination theory (SDT) provides the main framework for interpreting physical-activity motivation as a continuum from controlled to autonomous regulation. Across exercise and health research, autonomous regulation and satisfaction of basic psychological needs are consistently associated with more persistent activity patterns, whereas externally controlled regulation is less stable (Ryan & Deci, 2000; Teixeira et al., 2012; Ntoumanis et al., 2021; Ntoumanis & Moller, 2025). A recent narrative review similarly emphasizes internalized goals and autonomy as central to sustained fitness behavior (Gao et al., 2024).

University studies nevertheless differ substantially in instruments and outcomes. Sáez et al. (2021) linked motivational regulation, sport satisfaction, and activity in students, while Durán-Vinagre et al. (2023) examined BREQ-3-based motivational processes and intention to remain active. Vučković et al. (2022) combined exercise-motive and physical-activity measures in a post-lockdown student sample, and Yu et al. (2024) examined motivation, leisure satisfaction, mental toughness, and self-efficacy. These studies support the relevance of motivational regulation but also illustrate the dominance of cross-sectional self-report designs.

Social and environmental context adds another layer. Social support is associated with physical activity in university populations (Huang et al., 2025; Wang et al., 2024), while campus infrastructure and built-environment characteristics correlate with student health and fitness indicators (Ding et al., 2024; Lu et al., 2022). Broader evidence also shows the challenge of insufficient activity and sedentary behavior among young people and during periods of disrupted routines (Guthold et al., 2020; Hall et al., 2021).

Digital tools can support monitoring and engagement, but their effects depend on design and context. Wearable-tracker and mHealth studies report heterogeneous results across motivational interviewing, gamification, and community-support models (Nuss et al., 2023; Sañudo et al., 2024; Ash et al., 2025). Technology-supported physical-education designs also emphasize the importance of integrating digital tools with pedagogical structure rather than treating technology as an independent intervention (Sultoni et al., 2022).

In Kazakhstan, research on digital culture and technology-supported learning has highlighted changes in student engagement and the motivational potential of immersive tools (Tleshova & Gerfanova, 2024; Kussainova et al., 2025). Learning motivation has also been discussed as an indicator and condition of psychological well-being (Abildina et al., 2025). At the institutional level, the Health Promoting University framework integrates policy, infrastructure, programs, and participation, but implementation patterns remain context-dependent (Tafireyi & Grace, 2024; Comellas-Sáenz et al., 2025).

Prior studies therefore provide strong evidence for the relevance of autonomous motivation, social support, and environmental context, but they differ in design, measurement, and population. To clarify what

can and cannot be inferred across this literature, Table 1 summarizes representative studies that directly inform the present design.

Table 1

Selected empirical evidence relevant to motivational profiles, social support, learning context, and physical activity

Study / country-sample	Design & instruments	Digital tools	Statistics	Main finding	Limitation relevant here
Sáez et al., 2021 / Spain, n=1,099 students	Cross-sectional; short BREQ-2, PA hours, sport satisfaction	No	Correlations, t tests, CFA	Intrinsic regulation related to greater PA and satisfaction	Self-report; cross-sectional; two universities
Durán-Vinagre et al., 2023 / Spain, n=1,524 students	Cross-sectional; BREQ-3 and MIFAU	No	CFA, correlations, group comparisons	Self-determined regulation associated with intention to remain active	Cross-sectional; intention rather than objective PA
Nuss et al., 2023 / USA, n=40 inactive adults	12-week RCT; motivation and PA	Wearable tracker	Between-condition comparisons	Wearable effects depended on intervention combination	Small non-student sample
Sañudo et al., 2024 / Spain, 105 allocated young adults	8-week randomized mHealth trial; accelerometry and IPAQ	App + activity trackers	RCT group comparisons	Gamified mHealth improved selected PA outcomes	Young adults; attrition
Ash et al., 2025 / USA, 335 analytic	Retrospective longitudinal program analysis	MoveSpring + tracker/pedometer	χ^2 , survival/regression	Engagement and social features associated with retention/goals	Single university; mixed students/staff
Huang et al., 2025 / n=1,530 university students	Survey/path analysis; support, empowerment, exercise behavior	Not central	Path/mediation models	Support and empowerment associated with behavior-change processes	Context-specific; causal caution
Lu et al., 2022 / China, 4 colleges, n=1,498	Cross-sectional GIS + fitness records	GIS	Correlations	Built environment correlated with fitness outcomes	Fitness $\neq$ motivation; cross-sectional

The comparison shows three recurring limitations in the available evidence: heavy reliance on cross-sectional self-report designs, limited integration of motivational and contextual variables within the same university sample, and heterogeneous use of digital measures. The present study therefore focuses on descriptive and associational patterns in a Kazakhstani higher-education context and avoids causal claims that require longitudinal or experimental evidence.

Materials and research methods

The study was conducted at S. Seifullin Kazakh Agrotechnical Research University in Astana, Kazakhstan. The final analytical sample comprised 250 first- and second-year undergraduate students (185 men and 65 women) aged 17–20 years from the Institute of Energy, area of training 6B071 Engineering and Engineering Trades, educational program 6B07102 Electric Power Engineering. The sampled cohorts covered semesters 1–4. The physical-education component considered in the study was delivered over two semesters, with one scheduled class per week. Both first- and second-year students were represented; year-specific counts were not retained in the archived summary dataset.

Participants were recruited during regular university physical-education activities. Inclusion criteria were enrollment in the specified program and study years, participation in regular physical-education classes, and completion of the study questionnaire. Students with a formal medical exemption from practical physical education, an individualized or special study arrangement that prevented participation in regular classes, or incomplete study data were excluded before the final analysis. All 250 eligible students who were invited to

the final study sample participated; no refusals were recorded. The exact number excluded before invitation because of medical or special-study status was not retained. Participation was voluntary and anonymous.

A quantitative observational design was used to examine associations among motivational regulation, learning context, social support, and physical activity. The core study instrument was a 12-item author-developed questionnaire informed by self-determination theory and previous university physical-activity research. It comprised four blocks: autonomous regulation (4 items), controlled regulation (3 items), social support (3 items), and subjective appraisal of physical activity (2 items). All items were rated on a five-point Likert scale from 1 (strongly disagree) to 5 (strongly agree).

The autonomous-regulation items were: “I engage in physical activity because I enjoy it”; “Physical activity is a personal value for me”; “I view physical activity as a way of taking care of myself”; and “I feel an inner need to be physically active.” The controlled-regulation items were: “I engage in physical activity because teachers or people around me expect it”; “I engage in physical activity to avoid negative consequences (e.g., criticism or poor grades)”; and “Physical activity feels more like an obligation than a personal choice.”

The social-support items were: “My friends support me in being physically active”; “Teachers encourage me to be physically active”; and “It is easier for me to be physically active in a group or team.” The subjective-appraisal items were: “I am satisfied with my current level of physical activity” and “I believe physical activity positively affects my well-being and academic performance.” The questionnaire was conceptually informed by studies using established exercise-motivation instruments, including BREQ-2/BREQ-3, but the present 12-item form was not treated as a verbatim reproduction of a standardized scale.

Current-sample internal-consistency coefficients could not be recalculated from the archived aggregate dataset; therefore, no Cronbach’s alpha or McDonald’s omega values are retroactively reported. This limitation is stated explicitly below. The original analysis retained three descriptive motivational-profile categories (autonomous, controlled, and low motivation) based on the archived composite-score classifications. Because the participant-level file used to construct these categories was unavailable for the present revision, the manuscript no longer describes them as statistically validated or stable clusters and does not attribute an unverified Ward or k-means procedure to the original analysis.

Physical activity was summarized using the mean number of active days per week, and the archived summary also contained mean daily activity-time indicators for face-to-face, hybrid, and distance learning contexts. The original manuscript referred to consumer wearables and mobile applications, but device-level metadata, participant coverage, synchronization procedures, and missing-data logs were not retained. Accordingly, digital-device records are not treated as an independent empirical measure in the revised analysis; digital tools are discussed only as contextual evidence from prior research.

For the revised analysis, inferential statistics that could be reconstructed from reported aggregate values were calculated transparently. Differences in active days across the three motivational profiles were assessed with one-way ANOVA, η^2 , Bonferroni-adjusted pairwise comparisons, Cohen’s *d*, and confidence intervals. Pearson correlations between social-support indicators and physical activity were accompanied by 95% confidence intervals, *t* tests, and r^2 . Learning-format values are presented descriptively because the archived summary did not preserve format-specific sample sizes or whether observations were independent or repeated; inferential *p*-values and effect sizes are therefore not reported for that table.

Data collection and interpretation were organized in three stages. First, the diagnostic stage collected questionnaire responses and physical-activity indicators. Second, the analytical stage summarized motivational-profile categories, compared weekly activity across profiles, described activity indicators across learning formats, and examined correlations with social support. Third, the interpretive stage contextualized statistically observed associations; no causal effect of learning format, digital tools, or social support was inferred.

The study is therefore interpreted as a single-sample quantitative observational analysis rather than a mixed-methods or intervention study.

All statistical tests are two-sided. Exact *p*-values are reported where available; otherwise $p < .001$ is used. English decimal points are used consistently throughout the revised manuscript.

The analytic emphasis is on effect size, uncertainty, and association rather than on causal language.

Results and discussion

The revised results focus on the measures that have a visible analytical trace in the archived dataset: motivational-profile categories, weekly physical activity, descriptive learning-format indicators, and social-support correlations. Unsupported references to a separate qualitative strand, a standalone SDT scale, and independent digital-device effects have been removed.

The profile and social-support analyses are inferential where the necessary aggregate values are available. Learning-format statistics are presented descriptively because group-specific denominators and the exact repeated-versus-independent observation structure were not preserved.

1. Students' motivational profiles

The archived analysis classified students into three descriptive motivational profiles: autonomous motivation ($n = 95$; 38%), controlled motivation ($n = 110$; 44%), and low motivation ($n = 45$; 18%). These labels are retained as descriptive profile categories rather than statistically validated clusters. Weekly physical activity differed substantially across the three groups (Table 2).

Table 2

Motivational profiles and weekly physical activity ($N = 250$)

Motivational profile	n	%	Active days/week, M (SD)	95% CI for M
Autonomous motivation	95	38%	4.60 (1.20)	[4.36, 4.84]
Controlled motivation	110	44%	3.20 (1.00)	[3.01, 3.39]
Low motivation	45	18%	1.40 (0.70)	[1.19, 1.61]

One-way ANOVA based on the reported group sizes, means, and standard deviations showed a large between-profile difference, $F(2, 247) = 149.17$, $p < .001$, $\eta^2 = .547$. Bonferroni-adjusted pairwise comparisons were significant for all pairs: autonomous vs controlled, $\Delta M = 1.40$, Cohen's $d = 1.28$, familywise 95% CI [1.05, 1.75], $p < .001$; autonomous vs low motivation, $\Delta M = 3.20$, $d = 3.00$, 95% CI [2.75, 3.65], $p < .001$; and controlled vs low motivation, $\Delta M = 1.80$, $d = 1.95$, 95% CI [1.36, 2.24], $p < .001$. These differences describe the observed sample and do not establish a causal effect of motivational regulation.

2. Physical activity across learning formats

The archived summary contained physical-activity indicators for face-to-face, hybrid, and distance learning contexts. Because format-specific sample sizes and the observation structure were not retained, these values are reported descriptively rather than as an inferential comparison (Table 3).

Table 3

Descriptive physical-activity indicators by learning format

Learning format	Active days/week, M (SD)	Activity time, min/day, M (SD)
Face-to-face	3.80 (1.10)	52 (15)
Hybrid	3.50 (1.00)	49 (14)
Distance	2.60 (0.90)	37 (12)

Mean activity was higher in the face-to-face and hybrid contexts than in the distance context. However, the revised manuscript does not describe these differences as statistically significant because the archived summary does not support a valid reconstruction of the required group-specific inferential test. The values should therefore be interpreted as descriptive patterns only.

3. Social support and team-based participation

Pearson correlations were used to examine the association between social-support indicators and physical activity. The analysis covered support from friends, support from teachers, and participation in team-based activities. All three associations were positive (Table 4).

Table 4*Associations between social support and physical activity (N = 250)*

Indicator	r	95% CI	t(248)	p	r ²
Support from friends	.42	[.31, .52]	7.29	< .001	.176
Support from teacher	.35	[.24, .45]	5.88	< .001	.123
Team participation	.48	[.38, .57]	8.62	< .001	.230

The strongest association was observed for participation in team-based activities ($r = .48$), followed by support from friends ($r = .42$) and support from teachers ($r = .35$). The confidence intervals exclude zero, and all three tests were statistically significant at $p < .001$. Pearson's r and r^2 quantify association strength; p -values indicate evidence against the null hypothesis of zero correlation and should not be interpreted as measures of reliability or causality.

These findings are consistent with evidence that social support and empowerment are associated with university students' physical activity (Huang et al., 2025; Wang et al., 2024). The present results nevertheless remain correlational and do not show that social support independently causes higher activity.

Discussion

The main pattern is the substantial difference in weekly activity across the three motivational-profile categories. Students in the autonomous profile reported the highest activity, which is consistent with SDT-based evidence linking more self-determined regulation to persistent physical activity (Durán-Vinagre et al., 2023; Gao et al., 2024). Controlled motivation was associated with intermediate activity, while the low-motivation profile showed the lowest mean.

Learning-format values showed a descriptive gradient, with higher mean activity in face-to-face and hybrid contexts than in the distance context. Because the archived summary does not preserve the denominators required for a valid between-format test, this pattern is interpreted cautiously and is not presented as evidence that a learning format caused the observed differences.

Social support and team participation showed moderate positive correlations with physical activity. This is compatible with prior evidence emphasizing social interaction and support as motivational resources, but the observational design does not establish temporal direction or rule out confounding (Huang et al., 2025; Wang et al., 2024).

The digital literature provides useful context rather than an independent empirical finding in this study. Wearables, mHealth, and gamification can support activity monitoring and engagement, but effects vary by intervention design, social context, and participant group (Nuss et al., 2023; Sañudo et al., 2024; Ash et al., 2025). The revised manuscript therefore does not claim that digital tools strengthened motivation in the present sample.

In practical terms, the results support attention to autonomy-supportive teaching, team-based participation, and context-sensitive organization of physical education. These implications should be treated as targets for future intervention testing rather than as effects already demonstrated by the present observational data.

The single-university context is also important. Rather than generalizing to all Kazakhstani universities, the findings provide a focused description of first- and second-year engineering students within one institutional environment.

This interpretation is consistent with the Health Promoting University perspective, which treats student activity as embedded in institutional policy, infrastructure, and participation rather than as an individual motivational issue alone (Tafireyi & Grace, 2024; Comellas-Sáenz et al., 2025).

Overall, the evidence is strongest for between-profile differences in weekly activity and for positive associations between social support and activity. The learning-format pattern is descriptive, and no independent digital-tool effect is claimed.

The revised interpretation therefore emphasizes effect sizes, confidence intervals, and limits of inference

rather than causal wording. This aligns the conclusions with the actual observational structure of the study.

The convergence of motivational-profile differences and social-support correlations provides a coherent descriptive pattern, but psychometric and classification uncertainty must be considered because the archived aggregate dataset does not permit recalculation of internal-consistency or profile-stability statistics.

The results are broadly consistent with international studies of motivation, social support, and university physical activity, while the present contribution is the integration of these dimensions in a Kazakhstan-based engineering-student sample.

The findings have practical relevance for:

- physical-education instructors, by supporting the use of autonomy-supportive and team-based practices;
- university administrators, by highlighting the need to consider social support and learning context when planning student-activity initiatives;
- future research and program evaluation, by identifying motivational-profile and social-support indicators that can be tested in longitudinal or controlled designs.

These applications are recommendations for program design and evaluation; they should not be interpreted as causal effects demonstrated in the current dataset.

Limitations

This study has several limitations. First, recruitment was voluntary and non-probability based within a single university, so selection bias cannot be excluded and the findings should not be interpreted as representative of all university students in Kazakhstan. Detailed year-specific counts and the number excluded before invitation were not retained in the archived summary.

Second, the observational design does not establish temporal or causal relationships among motivational profile, learning format, social support, and physical activity. Academic workload, sex/gender, previous sport experience, baseline fitness, and access to sports infrastructure were not comprehensively controlled, and between-university variation could not be evaluated because the sample came from one institution.

Third, part of the evidence relies on self-report. Current-sample internal-consistency coefficients for the author-developed questionnaire could not be recalculated from the archived aggregate dataset, which limits psychometric interpretation of the block scores.

Fourth, device-level metadata for the digital records mentioned in the original manuscript were not available; consequently, the revised analysis does not treat heterogeneous consumer wearables or mobile applications as an independent empirical measure. Learning-format statistics are also descriptive because format-specific sample sizes and the observation structure were not preserved.

Finally, the motivational profiles are retained as descriptive categories from the archived analysis, not as statistically validated clusters. Future studies should retain participant-level data, use validated motivation instruments alongside transparent profile-classification or clustering procedures, standardize activity measurement, and apply longitudinal or controlled designs with multivariable adjustment.

These limitations define the scope of inference and provide a concrete agenda for replication and extension.

Conclusion

This study identified a coherent pattern linking motivational profile, social support, learning context, and physical activity among 250 first- and second-year engineering students at S. Seifullin Kazakh Agrotechnical Research University. The conclusions are observational and associational.

Autonomous-motivation profile students reported the highest weekly activity (4.60 ± 1.20 days), followed by controlled-motivation (3.20 ± 1.00) and low-motivation students (1.40 ± 0.70); the between-profile difference was large, $F(2, 247) = 149.17$, $p < .001$, $\eta^2 = .547$.

Face-to-face and hybrid learning contexts showed higher descriptive mean activity than the distance context, but the available archived summary does not support a valid causal or inferential between-format conclusion.

Support from friends, support from teachers, and team participation were positively associated with physical activity ($r = .35-.48$, all $p < .001$), with team participation showing the strongest correlation.

Taken together, the findings support the pedagogical relevance of autonomy-supportive and socially supportive physical-education environments while underscoring the need for longitudinal and controlled research before causal mechanisms are inferred.

The practical implications are:

to consider students' motivational profiles when planning differentiated physical-education and engagement strategies;

to strengthen peer, teacher, and team-based support as elements of university activity programs;

to evaluate learning-context and digital initiatives with standardized measures and prospective designs rather than assuming effectiveness from cross-sectional differences.

Future research should replicate the profile structure using retained participant-level data, validated psychometric instruments, standardized physical-activity measurement, and samples from multiple universities.

The revised evidence therefore identifies meaningful associations among motivation, social support, learning context, and physical activity while keeping the strength of the conclusions proportionate to the available data.

Conflict of Interest Statement

The authors declare no potential conflicts of interest regarding the research, authorship, or publication of this article.

Declaration of the use of AI

Generative AI tools were used solely for language editing, translation, and improvement of clarity and readability during manuscript revision. They were not used to generate, modify, or analyze empirical data; calculate statistical results; create tables or figures; formulate research findings; or generate references. All scientific content, data, analyses, tables, interpretations, and references were reviewed and approved by the authors, who remain fully responsible for the manuscript.

Author Contributions

All authors contributed equally to all stages of the research, including conceptualization, methodology, data analysis, manuscript preparation, review, and editing. All authors have read and approved the final version of the manuscript.

References

- Abildina, S., Omarova, M., & Kudarinova, A. (2025). Learning motivation as an indicator and condition of learners' psychological well-being. *Higher School of Kazakhstan*, 49(1), 105–114. <https://doi.org/10.59787/2413-5488-2025-49-1-105-114>
- Ash, G. I., Mak, S. S., Haughton, A. D., Augustine, M., Bodurtha, P. O., Axtell, R. S., Borsari, B., Liu, J. J., Lou, S., Xin, X., Fucito, L. M., Jeon, S., Stults-Kolehmainen, M., & Gerstein, M. B. (2025). College community-based physical activity support at a public university during the COVID-19 pandemic: Retrospective longitudinal analysis of intra- versus interpersonal components for uptake and outcome association. *JMIR mHealth and uHealth*, 13, e51707. <https://doi.org/10.2196/51707>
- Comellas-Sáenz, C., Esqué Boldú, E., & Juvinyà Canal, D. (2025). Health promoting universities: A scoping review. *Health Promotion Journal of Australia*, 36(2), e939. <https://doi.org/10.1002/hpja.939>
- Ding, Y., Lee, C., Chen, X., Song, Y., & Sohn, W. (2024). Exploring the association between campus environment of higher education and student health: A systematic review of findings and measures. *Urban Forestry & Urban Greening*, 91, 128168. <https://doi.org/10.1016/j.ufug.2023.128168>
- Durán-Vinagre, M. Á., Ibáñez, S. J., Feu, S., & Sánchez-Herrera, S. (2023). Analysis of the motivational processes involved in university students' physical activity. *Frontiers in Psychology*, 13, 1080162. <https://doi.org/10.3389/fpsyg.2022.1080162>
- Gao, Z., Chee, C. S., Omar Dev, R. D., & Gao, J. (2024). Comprehensive analysis of college students' autonomous fitness behavior—A narrative review. *Frontiers in Sports and Active Living*, 6, 1406810. <https://doi.org/10.3389/fspor.2024.1406810>

- Guthold, R., Stevens, G. A., Riley, L. M., & Bull, F. C. (2020). Global trends in insufficient physical activity among adolescents: A pooled analysis of 298 population-based surveys with 1.6 million participants. *The Lancet Child & Adolescent Health*, 4(1), 23–35. [https://doi.org/10.1016/S2352-4642\(19\)30323-2](https://doi.org/10.1016/S2352-4642(19)30323-2)
- Hall, G., Laddu, D. R., Phillips, S. A., Lavie, C. J., & Arena, R. (2021). A tale of two pandemics: How will COVID-19 and global trends in physical inactivity and sedentary behavior affect one another? *Progress in Cardiovascular Diseases*, 64, 108–110. <https://doi.org/10.1016/j.pcad.2020.04.005>
- Huang, J., Lv, Q., & Zeng, X. (2025). The influence of social support and empowerment on physical exercise behavior in university students: A self-determination theory perspective. *Acta Psychologica*, 257, 105086. <https://doi.org/10.1016/j.actpsy.2025.105086>
- Kussainova, A., Zhussupova, R., & Shadiev, R. (2025). Enhancing competencies and motivation: The impact of 360-degree video technology on future foreign language teachers. *Higher School of Kazakhstan*, 49(1), 67–82. <https://doi.org/10.59787/2413-5488-2025-49-1-67-82>
- Lu, Z., Li, Z., Mao, C., Tan, Y., Zhang, X., Zhang, L., Zhu, W., & Sun, Y. (2022). Correlation between campus-built environment and physical fitness in college students in Xi'an—A GIS approach. *International Journal of Environmental Research and Public Health*, 19(13), 7948. <https://doi.org/10.3390/ijerph19137948>
- Ntoumanis, N., & Moller, A. C. (2025). Self-determination theory informed research for promoting physical activity: Contributions, debates, and future directions. *Psychology of Sport and Exercise*, 80, 102879. <https://doi.org/10.1016/j.psychsport.2025.102879>
- Ntoumanis, N., Ng, J. Y. Y., Prestwich, A., Quested, E., Hancox, J. E., Thøgersen-Ntoumani, C., Deci, E. L., Ryan, R. M., Lonsdale, C., & Williams, G. C. (2021). A meta-analysis of self-determination theory-informed intervention studies in the health domain: Effects on motivation, health behavior, physical, and psychological health. *Health Psychology Review*, 15(2), 214–244. <https://doi.org/10.1080/17437199.2020.1718529>
- Nuss, K., Moore, K., Marchant, T., Courtney, J. B., Edwards, K., Sharp, J. L., Nelson, T. L., & Li, K. (2023). The combined effect of motivational interviewing and wearable fitness trackers on motivation and physical activity in inactive adults: A randomized controlled trial. *Journal of Sports Sciences*, 41(1), 45–55. <https://doi.org/10.1080/02640414.2023.2195228>
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Sáez, I., Solabarrieta, J., & Rubio, I. (2021). Motivation for physical activity in university students and its relation with gender, amount of activities, and sport satisfaction. *Sustainability*, 13(6), 3183. <https://doi.org/10.3390/su13063183>
- Sañudo, B., Sanchez-Trigo, H., Domínguez, R., Flores-Aguilar, G., Sánchez-Oliver, A., Moral, J. E., & Oviedo-Caro, M. Á. (2024). A randomized controlled mHealth trial that evaluates social comparison-oriented gamification to improve physical activity, sleep quantity, and quality of life in young adults. *Psychology of Sport and Exercise*, 72, 102590. <https://doi.org/10.1016/j.psychsport.2024.102590>
- Sultoni, K., Peralta, L. R., & Cotton, W. (2022). Using a design-based research approach to develop a technology-supported physical education course to increase the physical activity levels of university students: Study protocol paper. *PLOS ONE*, 17(12), e0269759. <https://doi.org/10.1371/journal.pone.0269759>
- Tafireyi, C. G. S., & Grace, J. M. (2024). Status of the health promoting university (HPU) globally and its relevance for emerging African HPUs: An integrative review and bibliometric analysis. *Global Health Promotion*, 31(3), 42–59. <https://doi.org/10.1177/17579759241235109>
- Teixeira, P. J., Carraça, E. V., Markland, D., Silva, M. N., & Ryan, R. M. (2012). Exercise, physical activity, and self-determination theory: A systematic review. *International Journal of Behavioral Nutrition and Physical Activity*, 9, 78. <https://doi.org/10.1186/1479-5868-9-78>
- Tleshova, Z., & Gerfanova, E. (2024). The Jean Monnet project research on digital culture and education: The analysis of categories in student responses. *Higher School of Kazakhstan*, 48(4), 135–147. <https://doi.org/10.59787/2413-5488-2024-48-4-135-147>
- Vučković, V., Krejač, K., & Kajtna, T. (2022). Exercise motives of college students after the COVID-19 lockdown. *International Journal of Environmental Research and Public Health*, 19(12), 6977. <https://doi.org/10.3390/ijerph19126977>

<https://doi.org/10.3390/ijerph19126977>

Wang, X., Yang, X., Mohd Nasiruddin, N. J. B., Wei, S., Dong, D., & Samsudin, S. B. (2024). Social support and physical activity in college and university students: A meta-analysis. *Health Education & Behavior*, 51(4), 533–543. <https://doi.org/10.1177/10901981231216735>

Yu, C., Zeng, Z., Xue, A., & Wu, Q. (2024). The effect of exercise motivation on college students' self-efficacy: The mediating roles of leisure satisfaction and mental toughness. *Frontiers in Psychology*, 15, 1465138. <https://doi.org/10.3389/fpsyg.2024.1465138>

Information about authors

Nurlan Mamiyev is a Candidate of Pedagogical Sciences and Senior Lecturer at the Department of Physical Education, S. Seifullin Kazakh Agrotechnical Research University, Astana, Kazakhstan. Email: n.mamiyev@kazatu.edu.kz ORCID: <https://orcid.org/0000-0001-7558-3353> (**corresponding author**)

Bakhytkhan Kuntugan is a Lecturer at the Department of Physical Education, S. Seifullin Kazakh Agrotechnical Research University, Astana, Kazakhstan. Email: b.kuntugan@kazatu.edu.kz ORCID: <https://orcid.org/0009-0004-5359-1227>

Arman Tukhmetov is a Senior Lecturer at the Department of Physical Education, S. Seifullin Kazakh Agrotechnical Research University, Astana, Kazakhstan. Email: a.tukhmetov@kazatu.edu.kz ORCID: <https://orcid.org/0009-0006-2080-1317>

Fakhrudin Batirov is a Senior Lecturer at the Department of Physical Education, S. Seifullin Kazakh Agrotechnical Research University, Astana, Kazakhstan. Email: f.batirov@kazatu.edu.kz ORCID: <https://orcid.org/0009-0004-5968-7203>

Received: 07.02.2026

Revised: 05.05.2026

Revised: 21.09.2026

Accepted: 30.09.2026

Published: 30.09.2026

Type of article: Research Article

¹Tannur Bakytказы, ²Kuralay Nurgaliyeva*, ³Eduarda Pinto Ferreira, ³Susana Nicola

¹ Abai Kazakh National Pedagogical University, Almaty, Kazakhstan

² Al-Farabi Kazakh National University, Almaty, Kazakhstan

³ Instituto Superior de Engenharia do Porto, Polytechnic of Porto, Porto, Portugal

DEVELOPING STEM COMPETENCIES IN PRE-SERVICE PHYSICS TEACHERS: A LONGITUDINAL UNIVERSITY-BASED INTERVENTION STUDY IN KAZAKHSTAN

Abstract. This multi-stage evaluation examines changes in physics teachers' self-assessed STEM competencies following a university-based professional development intervention in Kazakhstan. The study included a diagnostic needs survey (N = 164), course participation by 46 teachers, and complete paired pre-test/post-test data from 25 teachers, who also completed a practice-focused follow-up during the subsequent academic year. At baseline, 63.4% of respondents reported limited STEM understanding and 28.0% reported no familiarity. In the paired subsample, overall self-assessed STEM competence increased from $M = 2.18$ ($SD = 0.71$) to $M = 4.22$ ($SD = 0.58$), $t(24) = -13.78$, $p < .001$. Significant gains were observed in STEM knowledge, lesson planning, interdisciplinary task design, project-based learning, and digital-tool use. During follow-up, all 25 respondents reported implementing STEM projects, averaging 4.2 projects each; satisfaction was $M = 4.36$ ($SD = 0.70$), and 92% would recommend the course. These findings indicate substantial perceived professional growth and reported classroom transfer, but they do not establish causal effects because the study used convenience sampling, self-report measures, and no comparison group.

Keywords: STEM education, teacher competencies, physics teacher preparation, higher education, Kazakhstan

Introduction

STEM (Science, Technology, Engineering, and Mathematics) education has emerged as a strategic priority in contemporary global education systems, recognized by both researchers and policymakers as a key driver of innovation and economic development (Nazarbayev University Graduate School of Education, 2014; World Economic Forum, 2020). By fostering scientific literacy, technological fluency, critical thinking, creativity, and problem-solving skills, STEM-oriented approaches aim to prepare learners for increasingly complex knowledge-based societies (Sirajudin et al., 2021). Central to this paradigm is the integration of disciplinary knowledge with authentic real-world contexts, enabling the practical application of theoretical concepts (Bakytказы et al., 2024).

Within the STEM framework, physics occupies a foundational position, as it provides the conceptual basis for engineering design, technological systems, and scientific modeling (Balta et al., 2024). However, effective STEM integration in physics education depends substantially on teachers' professional competencies, including subject mastery, technological proficiency, and methodological readiness (Nurgaliyeva et al., 2023). Empirical evidence suggests that teachers' STEM competencies significantly influence student engagement, achievement, and career aspirations in STEM-related fields (Clark & Zhang, 2018; Shageeva et al., 2020).

International scholarship increasingly emphasizes multidimensional competency models in STEM teacher preparation (Kurup et al., 2019; Rose et al., 2015). Existing studies identify problem solving, collaboration, technological proficiency, organizational functioning, and integrated pedagogical design as relevant domains. However, subject-specific evidence concerning physics teachers in post-Soviet professional-development settings remains limited. This study addresses that contextual gap while recognizing that self-assessment alone cannot establish observable instructional performance.

In Kazakhstan, educational reforms prioritize teacher development and STEM-oriented pedagogy (Yakavets & Dzhadrina, 2014; Nazarbayev, 2018). However, practicing physics teachers continue to report methodological, organizational, and resource-related barriers to interdisciplinary STEM implementation. Evidence from school practice can therefore inform both in-service professional development and the later design of pre-service physics-teacher curricula.

Addressing this gap, the present study aims to develop and empirically evaluate a structured university-based professional development model for strengthening STEM competencies among physics teachers and to identify implications for pre-service teacher education.

Literature Review

Theoretical frameworks of STEM competencies. International scholarship proposes diverse frameworks for conceptualizing STEM competencies in pre-service teacher education. Berisha and Vula (2021) identify five pedagogical strategies: design-based, project-based, and problem-based learning, collaborative approaches, and inquiry or modeling, as central to competence development. Among these, design-based learning has been identified as particularly effective for interdisciplinary integration (Agudelo Rodríguez et al., 2024). However, while these models provide valuable pedagogical mechanisms, they often lack structured competency assessment frameworks applicable at the university level. Competency structure research further identifies four interrelated domains: problem-solving, collaboration, technological proficiency, and organizational functioning (Rose et al., 2015). Although such multidimensional models offer conceptual clarity, empirical validation in subject-specific teacher education contexts, particularly in physics, remains limited. Studies evaluating STEAM competencies frequently rely on self-assessment instruments and Likert-type scales (Smyrnova-Trybulska et al., 2024), which, while methodologically practical, may not fully capture observable instructional performance.

Pedagogical models and curriculum approaches. University-based teacher education programs increasingly combine multiple pedagogical models to foster STEM integration. Pimthong and Williams (2021) emphasize methodology courses as key platforms for developing integrated instructional design skills. Similarly, Li (2025) proposes innovative curriculum integration strategies through competitive pedagogical formats. Research on creative STEM teacher preparation underscores the importance of aligning theory, practice, and policy dimensions (Milner-Bolotin & Martinović, 2025). Nevertheless, many initiatives remain fragmented, focusing either on isolated instructional techniques or short-term interventions rather than longitudinal competency development. Virtual practicum experiences have also demonstrated effectiveness in enhancing instructional design skills (Wu et al., 2021). However, the scalability and contextual adaptability of such approaches in different higher education systems require further investigation.

Implementation strategies in higher education institutions. At the institutional level, STEM teacher preparation varies significantly in structure and delivery modes. Industry-based immersive placements positively influence professional intentions and instructional practices (Hurley et al., 2023). Comparative analyses based on TALIS-2018 data highlight structural differences in professional development systems across post-Soviet contexts (Chernobay & Tashibaeva, 2020). Recent mixed-methods studies demonstrate that structured STEM courses can enhance teacher self-efficacy (Zhumabay et al., 2024b). Despite these advances, empirical studies integrating theoretical modeling, structured intervention, and longitudinal evaluation within a single university-based framework remain scarce, particularly in Kazakhstan.

Research aim. This study aims to develop and empirically evaluate a structured university-based professional development model for strengthening STEM competencies among physics teachers.

Research objectives: (1) examine baseline levels of STEM awareness and professional needs among physics teachers; (2) develop a theoretically grounded intervention model; (3) evaluate changes in self-assessed STEM competencies from pre-test to post-test; and (4) examine teachers' implementation experience during the following academic year.

Research contribution. The study contributes a contextually adapted five-level model of STEM integration, a structured 80-hour professional-development course, verified pre-test/post-test evidence of change in self-assessed competencies, and follow-up evidence of reported classroom implementation in Kazakhstan. The contribution is empirical and contextual rather than a claim of an entirely original theory or research method.

Methodology

The empirical participants were in-service physics teachers working in secondary schools in Kazakhstan. Their professional needs and implementation experiences were examined to improve continuing professional development and to generate evidence that may inform future pre-service physics-teacher curricula. The empirical intervention itself did not involve pre-service teachers.

The study was situated within a teacher-development continuum linking initial preparation, early-career induction, and continuing professional development. Findings from practicing teachers were treated as indicators of current professional needs, not as direct evidence of the effectiveness of pre-service programmes.

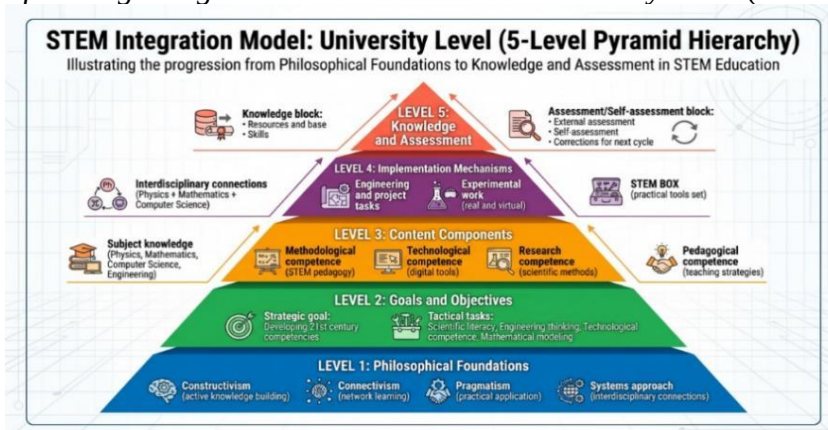
The five-level model was piloted as an organizing framework for the professional-development course. Its potential application in university teacher-education curricula remains a recommendation for future implementation and evaluation rather than a validated outcome of the present study.

The study examined whether a structured STEM course was associated with changes in teachers' self-assessed competencies and how participants reported applying STEM approaches during the following academic year. It did not independently measure student achievement. Accordingly, reported student-progress ratings are interpreted only as teacher perceptions, and no causal pathway from the intervention to student outcomes is claimed.

Research design. The study used a multi-stage quantitative evaluation supplemented by open-ended responses. First, 164 physics teachers completed a diagnostic needs survey. Second, 46 teachers participated in the university-based course "Integrating STEM Elements into Physics Teaching." Third, 25 teachers provided complete paired pre-test/post-test competency data and completed a practice-focused follow-up during the 2024–2025 academic year. The follow-up examined project implementation, barriers, perceived student progress, support, and satisfaction. Convenience sampling and the absence of a comparison group limit causal inference and generalizability.

Figure 1.

A Five-Level Model for Integrating STEM Education at the University Level (Penchev, 2025)



Data collection instruments. The diagnostic survey assessed STEM understanding, conditions required for effective lessons, missing competencies, difficult physics topics, preferred course content, and support needs. The intervention evaluation used five self-rating indicators on a 5-point scale: STEM knowledge, lesson planning, interdisciplinary task design, project-based learning, and digital-tool use. The follow-up questionnaire examined implementation frequency, project activity, barriers, perceived student progress, support, and satisfaction. The questionnaires also contained open-ended prompts about implemented projects and support needs.

The instruments were reviewed for alignment with the study's competency framework and piloted in the local context. Because complete item-level validation evidence and a fully specified measurement model were not available for this report, the five outcomes are treated as self-assessment indicators rather than validated latent subscales. The study therefore reports transparent indicator-level results and identifies further psychometric validation as a priority for future research.

Data analysis. Diagnostic responses were summarized using frequencies and percentages. Paired-samples t-tests were used for pre-test/post-test comparisons among the 25 participants with complete paired data; degrees of freedom were 24. Follow-up ratings are presented descriptively using means, standard deviations, frequencies, and percentages. Earlier omnibus ANOVA comparisons across difficulty categories and school subjects were removed because the same respondents rated multiple categories and participant-level data were unavailable for an appropriate repeated-measures or non-parametric reanalysis. Open-ended responses were used descriptively to identify examples of projects and support needs; no formal intercoder-

reliability claim is made.

The follow-up assessed reported transfer into classroom practice rather than administering the competency scale for a third time. The findings should therefore not be interpreted as a three-wave longitudinal test of sustained competency scores. Attrition is also relevant: 25 of the 46 course participants provided complete paired and follow-up information, and no systematic comparison with non-completers was available.

Ethical considerations. The study received approval from the Ethics Committee of Al-Farabi Kazakh National University. Participants provided informed consent, and confidentiality, anonymity, and the right to withdraw were maintained. The approval reference should be added in the journal submission metadata if required by the target journal.

Results

Teachers' understanding of and preparedness for STEM. The survey results (N=164) showed significant gaps in participants' understanding of STEM: STEM understanding level: – 63.4% (n=104) – limited understanding (general information only) – 28.0% (n=46) – not familiar at all – 8.5% (n=14) – in-depth, analytical-level knowledge. These findings indicate a considerable gap in teachers' knowledge and experience related to STEM education. Many teachers perceive STEM not as a comprehensive pedagogical system but as a set of isolated tasks and projects.

Factors necessary for effective implementation of STEM lessons (Figure 2) – Participation in professional development courses – 43.9% (n=72) – Collaboration and exchange of experience with colleagues – 26.8% (n=44) – Organization of interdisciplinary laboratories – 20.1% (n=33) – Partnership with universities – 6.7% (n=11) – Alignment of STEM activities with the school curriculum – 2.4% (n=4).

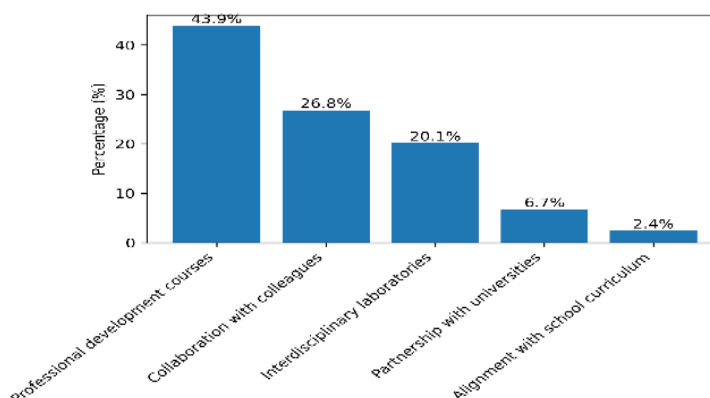
These indicators clearly demonstrate that teachers consider high-quality professional development courses to be the most critical need. In addition to improving their own skills, teachers emphasized the importance of institutional support, including adequate material and technical resources as well as organizational support.

In order to design the course, the competencies lacking for effective STEM integration were identified. These include insufficient methodological preparation – 24% (n=39); limited access to updated scientific information in the natural sciences – 24% (n=39); and effective organization of STEM education at the school level – 52% (n=86).

Difficult topics within the physics curriculum include Astronomy – 49.1% (n=81); Nuclear and Atomic Physics – 37.4% (n=61); Optics – 31.3% (n=51); Hydrodynamics and Wave Phenomena – 28.2% (n=46). These topics are generally considered challenging due to their complexity, lack of visual aids, and limited opportunities for practical activities. In response, the application of STEM strategies such as virtual laboratories (29.4%) and project-based learning (22.7%) can significantly enhance instructional effectiveness.

Figure 2

Factors Necessary for the Effective Implementation of STEM Lessons



Therefore, the course content should cover a broad range, from fundamental STEM concepts to more advanced methodologies and their practical application. Based on the survey data, the course was structured

around three main modules, as presented in Table 1.

To construct the course, an indicator-based analysis was conducted based on the survey responses. The results are presented in Table 1 below.

Table 1.
Indicator-Based Analysis Based on Survey Responses

Evidence area	Survey result	Priority need	Course response
STEM understanding	28.0% no familiarity; 63.4% limited; 8.5% in-depth	Foundational knowledge	Module 1
Missing preparation	50.0% organization; 24.4% methodology; 23.2% scientific updates	Organization and methodology	Modules 1 and 3
Teacher competencies	31.7% critical thinking; 28.7% lifelong learning; 25.0% organization; 22.6% leadership	Methodological development	Module 2
Challenging topics	49.1% astronomy; 37.4% nuclear physics; 31.3% optics; 28.2% waves	Visualization and practical work	Modules 2 and 3
Preferred approaches	29.4% virtual laboratories; 22.7% project learning; 17.8% applied tasks	Digital and project methods	Modules 2 and 3
Support needs	43.9% professional development; 26.8% peer exchange; 20.1% resources	Continuing professional support	All modules

Based on this analysis, the professional development course “Integrating STEM Elements into Physics Teaching” was developed. The course consisted of three main modules (Table 2).

Table 2
Structure and Content of the Professional Development Course

Module	Content
Module 1. Foundations of STEM Education (10 hours)	1.1. Regulatory documents 1.2. STEM: History and development 1.3. Principles of the STEM education system 1.4. STEM teaching practices in the USA 1.5. STEM integration in Kazakhstan: Opportunities and limitations 1.6. STEM teaching methodology: Critical and creative thinking
Module 2. Planning STEM Instruction (30 hours)	2.1. Core principles of STEM instructional planning 2.2. Approaches to applying STEM strategies in physics lessons 2.3. Developing project-based lesson plans 2.4. Specific features of STEM lessons for different grade levels 2.5. Methods of data collection and analysis 2.6. Reflection
Module 3. STEM Box and Resources: Project Development (30 hours)	3.1 STEM Box design and use; 3.2 efficient use of resources; 3.3 interdisciplinary project design; 3.4 project implementation; 3.5 presentation and defence; 3.6 assessment criteria. Final project presentation and assessment: 10 hours.

The course comprised 80 instructional hours: Module 1, Foundations of STEM Education (10 hours); Module 2, Planning STEM Instruction (30 hours); Module 3, STEM Box and Resources (30 hours); and final project presentation and assessment (10 hours). Instruction combined lectures and seminars with practical workshops, project development, virtual laboratories, webinars, Moodle, and Google Classroom. Participants completed a reflective task, a STEM lesson plan, an interdisciplinary project, and a final project presentation evaluated for curriculum alignment, interdisciplinary integration, feasibility, scientific accuracy, use of technology, and assessment design.

The five-level model was developed by adapting Penchev’s (2025) systems framework and related STEM-integration literature to the Kazakhstani professional-development context. It is presented as a

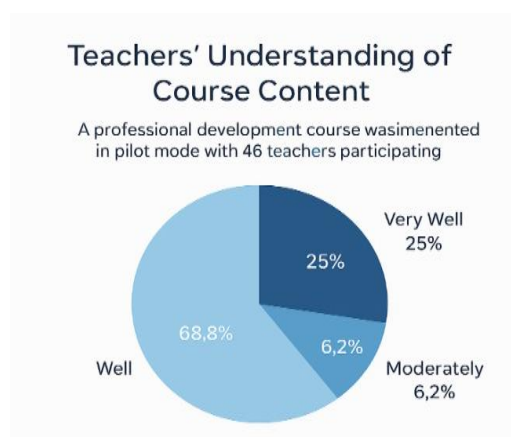
contextual synthesis rather than an entirely original or fully validated framework. The levels comprise philosophical foundations; goals and objectives; content components; implementation mechanisms; and education and assessment.

Results of course implementation. All 46 enrolled teachers completed the course, with an average attendance rate of 96%. Participants produced 46 STEM lesson plans, developed 23 interdisciplinary projects during the course, and participated in practical work with digital simulations, virtual laboratories, and project-based instructional design. The paired analysis reported below uses the 25 teachers for whom complete pre-test and post-test competency scores were available.

Immediate post-course feedback was positive. In the dissertation's course evaluation, perceived usefulness of the content was $M = 4.76$ ($SD = 0.44$), effectiveness of teaching methods was $M = 4.68$ ($SD = 0.48$), quality of learning materials was $M = 4.64$ ($SD = 0.49$), trainer professionalism was $M = 4.84$ ($SD = 0.37$), and overall satisfaction was $M = 4.72$ ($SD = 0.46$), all on a 5-point scale.

Figure 3

Immediate Post-Course Evaluation (Mean Ratings and Standard Deviations)



Participants reported that the course provided practical strategies for integrating STEM into physics instruction. Because these statements are based on self-report, they are interpreted as perceptions of readiness and usefulness rather than objective evidence of teaching mastery or student achievement.

Participants also perceived potential benefits for student engagement and interest in natural sciences. These perceptions were not verified with independent student-level measures and are therefore reported cautiously.

Paired pre-test/post-test comparisons showed statistically significant increases in all five indicators (Table 3). STEM knowledge increased from $M = 2.12$ ($SD = 0.83$) to $M = 4.28$ ($SD = 0.61$), $t(24) = -12.34$; lesson planning from $M = 1.88$ ($SD = 0.73$) to $M = 4.16$ ($SD = 0.69$), $t(24) = -13.45$; interdisciplinary task design from $M = 2.04$ ($SD = 0.79$) to $M = 4.08$ ($SD = 0.70$), $t(24) = -11.89$; project-based learning from $M = 2.32$ ($SD = 0.85$) to $M = 4.24$ ($SD = 0.66$), $t(24) = -10.67$; and digital-tool use from $M = 2.56$ ($SD = 0.92$) to $M = 4.36$ ($SD = 0.64$), $t(24) = -9.23$. Overall STEM competence increased from $M = 2.18$ ($SD = 0.71$) to $M = 4.22$ ($SD = 0.58$), $t(24) = -13.78$. All p values were $< .001$. Figure 4 visualizes the same means and standard deviations.

The follow-up evidence concerned classroom implementation during the subsequent academic year. All 25 respondents reported applying STEM projects, with an average of 4.2 projects per teacher. Reported projects addressed mechanics, electricity and magnetism, thermal physics, optics, and modern physics. The principal implementation difficulties are summarized in Table 4.

Table 3

Pre-Test and Post-Test STEM Competency Scores for Participants with Complete Paired Data (n = 25)

Competency indicator	Pre-test M (SD)	Post-test M (SD)	Mean change	t	p
STEM knowledge	2.12 (0.83)	4.28 (0.61)	+2.16	t(24) = -12.34	< .001
STEM lesson planning	1.88 (0.73)	4.16 (0.69)	+2.28	t(24) = -13.45	< .001
Interdisciplinary task design	2.04 (0.79)	4.08 (0.70)	+2.04	t(24) = -11.89	< .001
Project-based learning	2.32 (0.85)	4.24 (0.66)	+1.92	t(24) = -10.67	< .001
Digital-tool use	2.56 (0.92)	4.36 (0.64)	+1.80	t(24) = -9.23	< .001
Overall STEM competence	2.18 (0.71)	4.22 (0.58)	+2.04	t(24) = -13.78	< .001

Note. Scores were measured on a 5-point scale. Negative t values reflect the pre-test minus post-test direction. Participant-level data were unavailable for computing paired-difference confidence intervals or standardized effect sizes.

Table 4

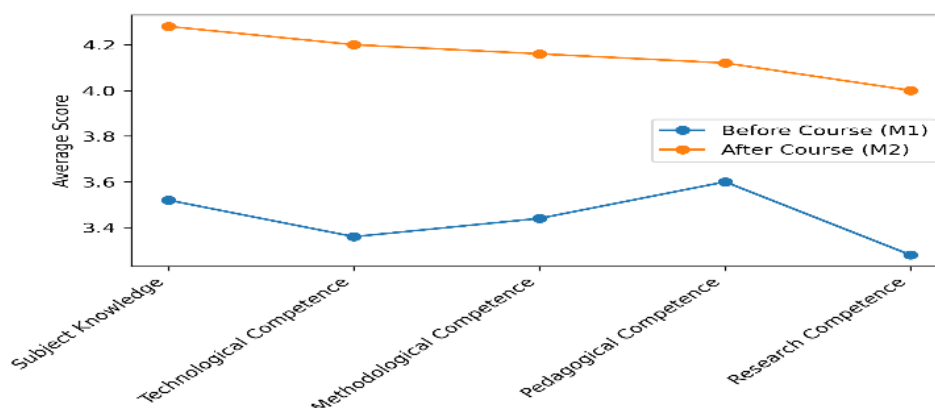
Difficulties Encountered During STEM Project Implementation (5-point scale, n = 25)

Difficulty	M	SD
Time constraints	2.57	1.42
Budget and funding	2.52	1.36
Student preparedness	2.48	1.23
Project assessment	2.44	1.23
Curriculum integration	2.36	1.29
Insufficient equipment	2.32	1.38
Interdisciplinary coordination	2.28	1.21
Insufficient methodological materials	2.24	1.33
Software availability	2.20	1.26
Parent support	2.12	1.17
Administrative support	2.08	1.19
Student grouping	2.00	1.15
Collegial support	1.96	1.10
Safety procedures	1.92	1.08
Internet access	1.87	1.19

Note. Ratings are presented descriptively because the same respondents evaluated all difficulty categories and participant-level data were unavailable for an appropriate repeated-measures analysis. The highest means concerned time constraints (M = 2.57), budget and funding (M = 2.52), student preparedness (M = 2.48), and project assessment (M = 2.44).

Figure 4

Pre-Test and Post-Test Self-Assessed STEM Competencies



During the course, trainees received various types of support in implementing STEM projects (Table 5).

Table 5*Types of Support Received and Their Effectiveness (5-point scale, N = 25)*

Type of support	Mean score	SD
Methodological support	4.24	0.72
Technical support	3.92	0.86
Collegial support	4.12	0.73
Administrative support	3.68	0.91
Financial support	3.44	1.02

At the one-year follow-up ($n = 25$), methodological support received the highest rating ($M = 4.24$, $SD = 0.72$), followed by collegial support ($M = 4.12$, $SD = 0.73$), technical support ($M = 3.92$, $SD = 0.86$), administrative support ($M = 3.68$, $SD = 0.91$), and financial support ($M = 3.44$, $SD = 1.02$). Overall satisfaction was $M = 4.36$ ($SD = 0.70$; 95% CI [4.09, 4.63]); 23 teachers (92%) would recommend the course, and all 25 reported applying acquired knowledge. Strengths included practical orientation (96%), professional support (88%), use of modern technologies (84%), and interdisciplinary integration (80%). Requested improvements included longer duration (52%), additional equipment (44%), and more methodological materials (36%). Perceived student-progress ratings ranged from $M = 3.64$ to 4.10 across six subjects and were highest in informatics ($M = 4.10$, $SD = 0.74$) and physics ($M = 4.00$, $SD = 0.71$). These descriptive ratings reflect teacher perceptions, not independently measured achievement.

Discussion

Baseline needs and intervention relevance. The diagnostic survey ($N = 164$) showed that 63.4% of respondents had only limited STEM understanding and 28.0% had no prior familiarity, while only 8.5% reported in-depth, practice-oriented understanding. Professional development was the most frequently selected condition for effective implementation (43.9%), followed by collaboration and experience exchange (26.8%) and interdisciplinary laboratories or material resources (20.1%). These findings directly informed the course's practical and methodological emphasis.

Intervention-associated change. Complete paired data from 25 teachers showed significant increases in all five self-assessment indicators and in the overall score (all $p < .001$). The largest mean increase occurred in lesson planning (+2.28), followed by STEM knowledge (+2.16), interdisciplinary task design (+2.04), project-based learning (+1.92), and digital-tool use (+1.80). These large changes warrant further confirmation with validated instruments, objective performance measures, and a comparison group.

Theoretical significance of the five-level model. The developed five-level model provides a comprehensive framework for integrating STEM education at the university level. The model synthesizes multiple theoretical approaches. Constructivism and connectivism (Level 1) ensure the principles of active

knowledge construction and networked learning (Suryadi et al., 2023), corresponding with Pimthong and Williams' (2021) concept of constructing integrated STEM education. The competence-oriented approach (Level 3) aligns with the multidimensional competence model proposed by Kurup et al. (2019), integrating subject, methodological, technological, research, and pedagogical competencies. The implementation mechanisms (Level 4) incorporate the five strategies proposed by Berisha and Vula (2021): design-based learning, project-based learning, problem-based learning, collaborative learning, and inquiry/modeling. The Knowledge and Assessment level (Level 5) applies the STEAM competence assessment approach of Smyrnova-Trybulska et al. (2024), combining self-assessment and external evaluation.

Practical implications. The findings suggest that teacher-education and professional-development providers may consider structured STEM components that emphasize practical work, interdisciplinary planning, digital tools, and continuing support. The present evidence does not justify mandatory national adoption without further multi-site evaluation. Extending support over 4–5 months was a participant recommendation rather than an experimentally tested conclusion.

Curriculum-level application. Based on the identified competency gaps and the outcomes of the intervention, a 5 ECTS elective course entitled “Elements of STEM Education in Physics Teaching” was developed for master’s students enrolled in the 7M01501 Physics educational programme. The purpose of the course is to develop students’ professional competencies through the study of the theoretical foundations, core principles, and methodological approaches of STEM education. The course includes project-based approaches to planning and conducting physics lessons and elective courses, methods for designing STEM boxes, and strategies for the effective use of educational resources. The development of this course represents a practical curriculum-level outcome of the study. However, its effectiveness among master’s students has not yet been empirically evaluated; therefore, its implementation and assessment constitute a direction for future research.

School support system. In accordance with the recommendations of Chernobay and Tashibaeva (2020), a comprehensive support system for schools should be established, including provision of technical equipment, a repository of methodological materials, partnership networks (inter-school exchange of experience), and administrative and financial support. Policy-level changes. Consistent with Zhumabay et al. (2024a), STEM education should be integrated into national educational standards, teachers’ workload should be optimized, funding for STEM resources should be increased, and partnerships between universities and schools should be strengthened.

The study has several limitations. Convenience sampling restricts generalizability, and the absence of a comparison group prevents causal attribution. Only 25 of 46 course participants provided complete paired and follow-up data, creating possible attrition bias. Outcomes were predominantly self-reported, and complete psychometric evidence for the indicators was not available. The follow-up did not repeat the competency scale, qualitative responses were not analyzed with independent coders, and student progress was not measured objectively. Difficulty and subject ratings are therefore reported descriptively rather than with an inappropriate independent-groups ANOVA. Future studies should use validated instruments, matched or randomized comparison groups, repeated-measures or non-parametric analyses appropriate to within-participant data, classroom observations, portfolio assessment, objective student outcomes, and transparent preregistered analysis plans.

Conclusion

This study found substantial pre-test/post-test increases in physics teachers’ self-assessed STEM competencies following an 80-hour, practice-oriented professional-development course. Overall competence increased from $M = 2.18$ to $M = 4.22$, $t(24) = -13.78$, $p < .001$. All 25 follow-up respondents reported implementing STEM projects during the subsequent academic year, while methodological and collegial support received the strongest ratings. These findings indicate perceived professional growth and reported classroom transfer, but they do not establish causal effectiveness or objectively verified improvement.

As a curriculum-level application of the findings, a 5 ECTS elective course entitled “Elements of STEM Education in Physics Teaching” was developed for master’s students in the 7M01501 Physics educational programme.

The study offers a contextually adapted framework and preliminary evidence-based recommendations for STEM teacher professional development in Kazakhstan. Further validation is required before broader

claims about effectiveness, student achievement, or transferability can be made.

Future research should extend this work through comparative studies across contexts, qualitative investigations of teacher experiences, and longitudinal tracking of career development. By continuing to build the empirical foundation for STEM teacher education, the research community can support evidence-informed policy and practice that strengthens STEM education systems globally.

Funding information

This research was carried out within the framework of the project AP26101969 “Development and Implementation of a Synergistic Model for the Training and Professional Development of Innovative and Active Teachers in the Context of Globalization, Digitalization, and Inclusive Education,” funded under the Grant Funding of Scientific Research program of the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan.

Conflict of Interest Statement

The authors declare no potential conflicts of interest regarding the research, authorship, or publication of this article.

Declaration of the use of AI

The authors declare that artificial intelligence (AI) tools were not used in the preparation of the manuscript text, data analysis, interpretation of the results, or formulation of the conclusions. AI tools were used solely for the generation of one or more visual materials/figures presented in the manuscript. The authors reviewed and verified the generated visual content and assume full responsibility for its accuracy, relevance, and correspondence to the scientific content of the article.

Acknowledgments

The authors express their sincere gratitude to all teachers who participated in the study, as well as to the Center for Natural Science Education of Al-Farabi Kazakh National University for their support.

Author Contributions

All authors contributed equally to all stages of the research, including conceptualization, methodology, data analysis, manuscript preparation, review, and editing. All authors have read and approved the final version of the manuscript.

References

- Agudelo Rodríguez, C. M., González-Reyes, R. A., Bernal-Ballén, A., et al. (2024). Characterization of STEM teacher education programs for disciplinary integration: A systematic review. *Eurasia Journal of Mathematics, Science and Technology Education*, 20(3). <https://doi.org/10.29333/ejmste/14280>
- Bakytказы, T., Nurgaliyeva, K., Japashov, N., & Zhumabay, N. (2024). Examining the pedagogical abilities and needs of Kazakh physics teachers to implement STEM education. In *Proceedings of the 27th International Conference on Interactive Collaborative Learning (ICL2024)* (pp. 1033–1044).
- Bakytказы, T., Nurgaliyeva, K., Japashov, N., & Zhumabay, N. (2024). Purpose and characteristics of STEM education in both the United States and Kazakhstan. In *Proceedings of the 27th International Conference on Interactive Collaborative Learning (ICL2024)* (pp. 1539–1546).
- Balta, N., Japashov, N., Bakytказы, T., Abdikadyr, B., & Nurgaliyeva, K. (2024). High school physics teachers' pedagogical discontentment: The effect of curriculum improvement and professional development programs. *Heliyon*, 10(2).
- Berisha, F., & Vula, E. (2021). Developing pre-service teachers' conceptualization of STEM and STEM pedagogical practices. *Frontiers in Education*, 6. <https://doi.org/10.3389/feduc.2021.585075>
- Carnegie, A. P., Smith, N., & Melton, M. (2011). *STEM: Science, technology, engineering, mathematics*. Georgetown University Center on Education and the Workforce.
- Chernobay, E. V., & Tashibaeva, D. N. (2020). Professional development of teachers in the Russian Federation and the Republic of Kazakhstan according to the results of the TALIS-2018 study. *Educational Issues*

- (Voprosy obrazovaniya), 4, 141–164.
- Clark, P. A., & Zhang, M. (2018). The impact of teacher technology training on student technology assessments. *Journal of Research in Science, Mathematics and Technology Education*, 1(3), 1–38.
- Golzar, J., Noor, S., & Tajik, O. (2022). Convenience sampling. *International Journal of Education & Language Studies*, 1(2), 72–77.
- Hurley, M., Healy, D., Hamilton, C., Butler, M., & McLoughlin, E. (2023). STEM teacher professional learning through immersive STEM learning placements in industry: A systematic literature review. *Journal for STEM Education Research*. <https://doi.org/10.1007/s41979-023-00089-7>
- Kurup, P. M., Li, X., Powell, G., & Brown, M. (2019). Building future primary teachers' capacity in STEM: Based on a platform of beliefs, understandings and intentions. *International Journal of STEM Education*, 6, 10. <https://doi.org/10.1186/s40594-019-0164-5>
- Li, T. (2025). Research on the integration model of STEM curriculum and teaching skills innovation competition for pre-service teachers: A new pathway for cultivating science education professionals. *International Journal of Education and Humanities*. <https://doi.org/10.54097/dbhzbkb31>
- Milner-Bolotin, M., & Martinović, D. (2025). Creative approaches for 21st-century science, technology, engineering, and mathematics teacher education: From theory to practice to policy. *Future in Educational Research*. <https://doi.org/10.1002/fer3.70>
- Nazarbayev, N. A. (2018). Growing prosperity, improving welfare of Kazakhstanis: New opportunities under the Fourth Industrial Revolution. State of the Nation Address.
- Nazarbayev University Graduate School of Education. (2014). Development of strategic directions for education reforms in Kazakhstan for 2015–2020: Diagnostic report. Indigo Print.
- Nurgaliyeva, K., Mirkhamitova, K., Gabdullina, A., & Igenbayeva, A. (2023). Analysis of students' satisfaction with distance learning and the advantages and disadvantages of laboratory works with computer simulation of electronic analog circuits. In *Proceedings of the 18th Conference on Electrical Machines, Drives and Power Systems (ELMA 2023)*.
- Penchev, D. (2025). STEM Integration as a System. *European Journal of STEM Education*, 10(1), 20. <https://doi.org/10.20897/ejsteme/17214>
- Pimthong, P., & Williams, J. (2018). Preservice teachers' understanding of STEM education. *Kasetsart Journal of Social Sciences*. <https://doi.org/10.1016/j.kjss.2018.07.017>
- Pimthong, P., & Williams, P. J. (2021). Methods course for primary level STEM preservice teachers: Constructing integrated STEM teaching. *Eurasia Journal of Mathematics, Science and Technology Education*, 17(8). <https://doi.org/10.29333/ejmste/11113>
- Polgampala, A. S. V., Shen, H., & Huang, F. (2017). STEM teacher education and professional development and training: Challenges and trends. *American Journal of Applied Psychology*, 6(5), 93–97.
- Rose, M. A., Shumway, S., Carter, V., & Brown, J. (2015). Identifying characteristics of technology and engineering teachers striving for excellence using a modified Delphi. *Journal of Technology Education*, 26(2).
- Shageeva, F. T., Mishchenko, E. S., Chernyshov, N. G., Turekhanova, K. M., & Omirzhanov, Y. T. (2020). International ENTER project: A new pedagogical training approach for engineering educators. *Vyshee Obrazovanie v Rossii*, 29(6), 65–74.
- Sirajudin, N., et al. (2021). Developing creativity through STEM education. *Journal of Physics: Conference Series*, 1806(1), 012211.
- Smyrnowa-Trybulska, E., Kopczyński, T., Morze, N., et al. (2024). Assessment and development of STEAM competences of pre-service and in-service teachers: Selected aspects. *Kultura i Edukacja*. <https://doi.org/10.15804/kie.2024.04.12>
- Suryadi, A., Purwaningsih, E., & Yuliati, L. (2023). STEM teacher professional development in pre-service teacher education: A literature review. *Waikato Journal of Education*, 28(1). <https://doi.org/10.15663/wje.v28i1.1063>
- World Economic Forum. (2020). The future of jobs report 2020. <https://www.weforum.org/reports/the-future-of-jobs-report-2020>
- Wu, B., Peng, X., & Hu, Y. (2021). How to foster pre-service teachers' STEM learning design expertise through virtual internship: A design-based research. *Educational Technology Research and*

- Development, 69, 3517–3543. <https://doi.org/10.1007/s11423-021-10063-y>
- Yakavets, N., & Dzhadrina, M. (2014). Educational reform in Kazakhstan: Entering the world arena. In D. Bridges (Ed.), Educational reform and internationalisation: The case of school reform in Kazakhstan (pp. 28–52). Cambridge University Press.
- Yip, W. Y. V. (2020). Developing undergraduate student teachers' competence in integrative STEM teaching. *Frontiers in Education*, 5, 44. <https://doi.org/10.3389/educ.2020.00044>
- Zhu, J. (2023). STEM pre-service teacher education: A review of research trends in the past ten years. *Eurasia Journal of Mathematics, Science and Technology Education*, 19(7). <https://doi.org/10.29333/ejmste/13300>
- Zhumabay, N., Varis, S., Abylkassymova, A., Bakytказы, T., & Bowen, G. M. (2024a). Mapping the Kazakhstani STEM education landscape: A review of national research. *European Journal of STEM Education*, 9(1), 16.
- Zhumabay, N., Yelemessova, Z., Balta, N., Abylkassymova, A., Bakytказы, T., & Marynowski, R. (2024b). Designing effective STEM courses: A mixed-methods study of the impact of a STEM education course on teachers' self-efficacy and course experiences. *Frontiers in Education*, 9, 1276828.

Information about authors:

Bakytказы Tannur - doctoral student, Abai Kazakh National Pedagogical University, Almaty, Kazakhstan, tannur1118@gmail.com, ORCID: 0009-0001-8606-2194

Nurgaliyeva Kuralay - Candidate of Physical and Mathematical Sciences, Assistant Professor, Al-Farabi Kazakh National University, Almaty, Kazakhstan, knurgaliyeva@kaznu.kz, ORCID: 0000-0002-0696-7277 (**corresponding author**)

Eduarda Pinto Ferreira - Professor, Instituto Superior de Engenharia do Porto, Polytechnic of Porto, Porto, Portugal, epf@isep.ipp.pt, ORCID: 0000-0002-6270-0382

Susana Nicola - Adjunct Professor, PhD in Industrial Engineering and Management, Instituto Superior de Engenharia do Porto, Polytechnic of Porto, Porto, Portugal, sca@isep.ipp.pt, ORCID: 0009-0005-3915-7421

Received: 27.02.2026

Revised: 14.06.2026

Revised: 31.08.2026

Accepted: 22.09.2026

Published: 30.09.2026

Type of article: Research Article

Aida Sagintayeva*, Gulmira Qanay
Nazarbayev University

BEYOND REGULATION: INSTITUTIONAL CAPACITY AND DOCTORAL EDUCATION IN KAZAKHSTAN

Abstract. Doctoral education has expanded across many higher education systems as governments seek to strengthen national research capacity, innovation, and international competitiveness. Yet growth in doctoral enrolment does not necessarily translate into stronger doctoral outcomes. This study examines recent developments in doctoral education in Kazakhstan through an institutional capacity perspective. Drawing on official statistics for 2021–2024 and international scholarship, it analyses trends in doctoral admissions, graduation, and completion with dissertation defence. Institutional capacity is not measured directly; rather, it is used as an interpretive framework centred on five organisational domains: supervision, research environments, governance, progress monitoring, and research integrity. The findings point not to a simple deterioration in doctoral outcomes, but to an uneven pattern of system development. Admissions remained relatively stable between 2021 and 2023 before increasing sharply in 2024, while graduation and dissertation defence followed different trajectories. These trends do not establish causal relationships but raise questions about whether organisational capacity develops alongside system expansion. The article distinguishes between expansion of PhD provision and the institutional development required to sustain it, arguing that regulatory reform and enrolment growth should be accompanied by greater attention to universities' capacity to support high-quality PhD provision.

Keywords: PhD education; institutional capacity; doctoral outcomes; supervision; research environment; higher education policy; Kazakhstan

Introduction

Doctoral education has become a strategic priority for higher education systems seeking to strengthen research capacity, innovation, and international competitiveness. Expansion, however, does not necessarily produce stronger doctoral outcomes. Universities operating within the same national regulatory framework may differ in progression, completion, research productivity, and doctoral experience, suggesting that national policy alone cannot fully explain how doctoral programmes perform. Governments establish regulatory and funding frameworks, while universities create the organisational conditions within which doctoral research is undertaken (McAlpine & Norton, 2006; Skopek et al., 2022; Taylor, 2023).

International scholarship increasingly conceptualises doctoral education as an institutional process rather than solely an individual learning experience or supervisory relationship. From this perspective, doctoral development is shaped by supervision, research environments, governance, progress monitoring, and research integrity, as well as formal programme requirements (Lee, 2008; Halse & Malfroy, 2010; EUA Council for Doctoral Education, 2016).

Kazakhstan provides a useful context in which to examine these issues. Following its adoption of the Bologna three-cycle degree structure, the country introduced substantial reforms intended to modernise PhD training, strengthen research capacity, and increase international integration (Ibraev et al., 2015; Kuzhabekova, 2020). Recent national statistics reveal an increasingly complex pattern of system development. Admissions remained broadly stable between 2021 and 2023 before rising sharply in 2024, while graduation and dissertation defence followed different trajectories (National Academy of Sciences of the Republic of Kazakhstan under the President of the Republic of Kazakhstan, 2025). These contrasting trends raise questions about how system-level expansion relates to the organisational conditions through which universities support PhD provision.

Existing research on PhD education in Kazakhstan has examined individual aspects of this transformation, including publication requirements, supervision, internationalisation, and research integrity, but less attention has been paid to how these organisational dimensions collectively relate to system-level change.

This article addresses that gap through an institutional capacity perspective. Institutional capacity is understood as universities' ability to create, coordinate, and sustain the organisational conditions required to support doctoral research and progression. It is not treated as a directly measured variable. Rather, it provides a framework for interpreting national trends through five interrelated domains: supervision, research environments, governance, progress monitoring, and research integrity.

The study makes three contributions. First, it integrates these organisational dimensions within a single institutional framework. Second, it demonstrates how national doctoral statistics can be interpreted through contemporary scholarship without treating aggregate trends as evidence of causal institutional effects. Third, it distinguishes system expansion from the institutional development required to sustain that expansion.

The study addresses the following questions:

How did doctoral admissions, doctoral completion within the normative period, and graduation with dissertation defence change in Kazakhstan between 2021 and 2024?

How can these trends be interpreted through an institutional capacity perspective?

What implications does this interpretation have for doctoral education policy in Kazakhstan and other rapidly expanding systems?

Literature Review

Doctoral education beyond regulation

Research on doctoral education has increasingly shifted from an emphasis on individual candidates and supervisory relationships towards the broader organisational environments in which doctoral research occurs (Lee, 2008; Halse & Malfroy, 2010; Taylor, 2023). The Salzburg Principles similarly position doctoral education as part of universities' research mission and emphasise strong research environments, critical mass, transparent supervision, and institutional responsibility (EUA Council for Doctoral Education, 2016).

This institutional turn is particularly relevant in systems governed by common national regulations. Formal requirements may be similar across universities, yet institutions differ in research intensity, supervisory capacity, academic culture, organisational support, and infrastructure. Evidence on doctoral completion also suggests that programme organisation and institutional factors matter for progression and time to completion (Skopek et al., 2022). National policy therefore establishes the formal framework for doctoral provision, but universities mediate how that framework operates in practice.

The distinction is also visible in other rapidly reforming post-Soviet systems. Ruziev et al. (2026), for example, show how doctoral reforms in Uzbekistan interact with institutional capacity constraints, incentives, and organisational practices. Such evidence reinforces the need to examine not only what regulations require, but also whether institutions possess the capacity to implement and sustain those requirements.

Organisational conditions supporting doctoral education

Supervision remains central to doctoral education, but it is increasingly understood as institutional and professional work rather than solely an interpersonal relationship. Universities shape supervisory practice through workload allocation, supervisor preparation, mentoring, co-supervision, monitoring, and quality assurance (Halse & Malfroy, 2010; McAlpine, 2013). Recent international research also emphasises the multiple academic, developmental, organisational, and relational functions of doctoral supervisors (Guarimata-Salinas et al., 2024).

Doctoral development also depends on the wider research environment. Research groups, seminars, conferences, collaborative projects, and scholarly networks provide opportunities for intellectual exchange and academic integration (McAlpine & Norton, 2006). Support from supervisors and research communities has been associated with doctoral candidates' research engagement (Rönkkönen et al., 2024). The Salzburg framework captures this through the concept of critical mass: sufficient concentrations of expertise, interaction, and resources to sustain doctoral research (EUA Council for Doctoral Education, 2016).

Governance and progress monitoring constitute a further organisational dimension. Doctoral programmes require coordination among supervisors, departments, doctoral committees, graduate schools, and institutional leadership. Annual reviews, milestone assessments, and early intervention mechanisms can shift responsibility for progression from an exclusively supervisor–candidate relationship towards a broader institutional process (McAlpine, 2013; Skopek et al., 2022).

Research integrity is similarly an institutional as well as individual responsibility. Universities influence

responsible research through ethics education, publication support, research governance, assessment procedures, and academic cultures. Research-integrity education for doctoral researchers is increasingly recognised as an institutional responsibility (Abdi et al., 2021). This is especially relevant where publication requirements strongly shape doctoral progression and academic incentives.

Together, these dimensions suggest that doctoral education operates through interconnected organisational conditions. Their interaction provides the basis for the institutional capacity perspective developed in this article.

Institutional capacity and doctoral education

Institutional capacity is defined here as a university's organisational ability to create, coordinate, and sustain the conditions necessary for effective doctoral training and research. It refers not simply to formal structures but to the capacity to translate national policy and programme requirements into functioning systems of academic support (McAlpine & Norton, 2006; McAlpine, 2013; Taylor, 2023).

This distinction matters because regulatory reform and institutional development may proceed at different speeds. Governments can expand doctoral places or revise degree and publication requirements relatively quickly, whereas developing experienced supervisory communities, strong research cultures, effective governance, and reliable support systems is cumulative and may require considerably more time (Skopek et al., 2022; Ruziev et al., 2026).

The five dimensions used in this study are supervision, research environments, governance, progress monitoring, and research integrity. They represent domains through which institutional capacity becomes relevant to doctoral provision. These dimensions are not measured directly. Instead, they provide a theoretically informed framework for interpreting national trends and identifying questions that aggregate statistics cannot answer.

Doctoral education in Kazakhstan

Kazakhstan has undertaken substantial reform of doctoral education following adoption of the Bologna Process. Reforms consolidated the three-cycle degree structure, strengthened quality assurance, revised publication requirements, and expanded international engagement (Ibraev et al., 2015; Kuzhabekova, 2020).

Recent scholarship demonstrates that these reforms interact with institutional conditions. Kuzhabekova (2025) and Bayanbayeva (2026) examine the consequences of publication requirements for research productivity and academic practice. Tokatligil et al. (2025) show that international co-supervision depends on effective institutional support and coordination. Satpayeva et al. (2025) identify continuing challenges in doctoral training in the social sciences and humanities, while Shamatov et al. (2026) examine research integrity under publication requirements.

These studies have advanced understanding of particular dimensions of doctoral education, but they have generally examined them separately. The present study brings these organisational dimensions together within an institutional capacity framework and uses that framework to interpret national trends during a period of rapid expansion.

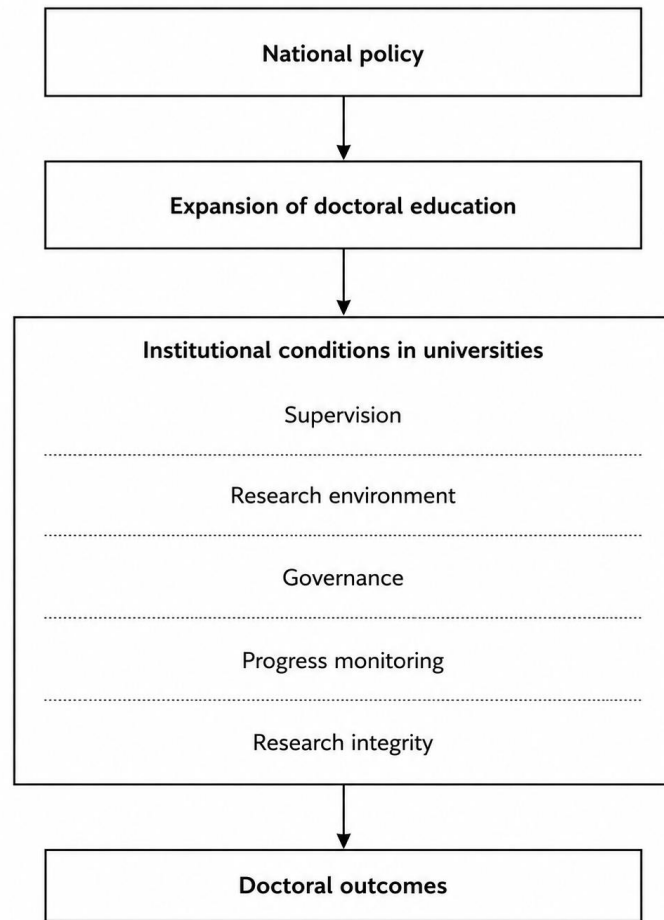
Conceptual Framework

The study is guided by the premise that national regulation alone cannot fully explain the development of doctoral education. Government policy establishes formal conditions through regulation, funding, quality assurance, admissions requirements, and research expectations. Universities translate these conditions into practice through their organisational structures, academic practices, and research environments (McAlpine & Norton, 2006; McAlpine, 2013; Taylor, 2023).

Figure 1 conceptualises this relationship across three elements: national policy context, institutional capacity, and doctoral outcomes. Institutional capacity comprises five interrelated domains (supervision, research environments, governance, progress monitoring, and research integrity) and represents the organisational conditions through which policy is translated into doctoral practice.

Figure 1.

Institutional Capacity and Doctoral Education: A Conceptual Framework



Note. The framework illustrates the relationship between the national policy context, five interrelated domains of institutional capacity, and doctoral outcomes. The relationship between institutional capacity and outcomes is theoretically informed and is not empirically tested in this study.

Source. Authors' conceptualisation informed by McAlpine and Norton (2006), McAlpine (2013), EUA Council for Doctoral Education (2016), Taylor (2023), and Skopek et al. (2022).

The framework does not assume a direct causal relationship between institutional capacity and the national indicators analysed here. The available data are aggregated and contain no university-level measures of the five domains. Institutional capacity therefore functions as an interpretive framework rather than an explanatory variable in the statistical sense.

Methodology

Research design and data sources

This study employs descriptive secondary-data analysis combined with theoretically informed interpretation. It examines national trends in doctoral education in Kazakhstan between 2021 and 2024 and interprets them through the institutional capacity framework developed in the preceding sections.

The empirical analysis draws on official national statistics reported by the National Academy of Sciences of the Republic of Kazakhstan (2025), based on Bureau of National Statistics data. Three indicators are examined: doctoral admissions, total doctoral graduates, and graduates completing with dissertation defence. These indicators provide a system-level picture of changes in doctoral provision and outcomes but do not include institution-level information on supervision, research environments, governance, progress monitoring, or research integrity.

International and Kazakhstan-focused scholarship is therefore used to provide the conceptual basis for

interpreting the observed trends. The study does not seek to establish causal relationships between institutional conditions and doctoral outcomes.

Data analysis

The analysis proceeded in two stages. First, the three national indicators were examined descriptively to identify changes and annual patterns between 2021 and 2024. Second, these patterns were interpreted through the five domains of institutional capacity identified in the literature: supervision, research environments, governance, progress monitoring, and research integrity. For each year, the proportion of graduates completing with dissertation defence was calculated by dividing the number of graduates with dissertation defence by the total number of doctoral graduates.

The indicators are treated as annual system-level measures rather than as a longitudinal cohort series. Admissions and graduation figures relate to different groups of doctoral candidates, while graduates completing with dissertation defence constitute a subset of total graduates in each year. Consequently, comparisons across these indicators are used to describe system development rather than to calculate cohort completion rates.

Institutional capacity is likewise not inferred directly from the statistical data. Rather, the observed trends identify issues requiring further explanation, while the conceptual framework provides theoretically grounded organisational dimensions through which those issues can be examined.

Limitations

The study has two principal limitations. First, aggregated national statistics cannot reveal variation across universities, disciplines, or doctoral programmes, nor can they track individual cohorts over time. Second, the available data do not directly measure the five domains of institutional capacity. The framework is therefore used for interpretation rather than causal explanation.

Future research should test the proposed relationships using institution-level data, comparative case studies, qualitative evidence from doctoral candidates and supervisors, and longitudinal analysis of doctoral progression.

Results

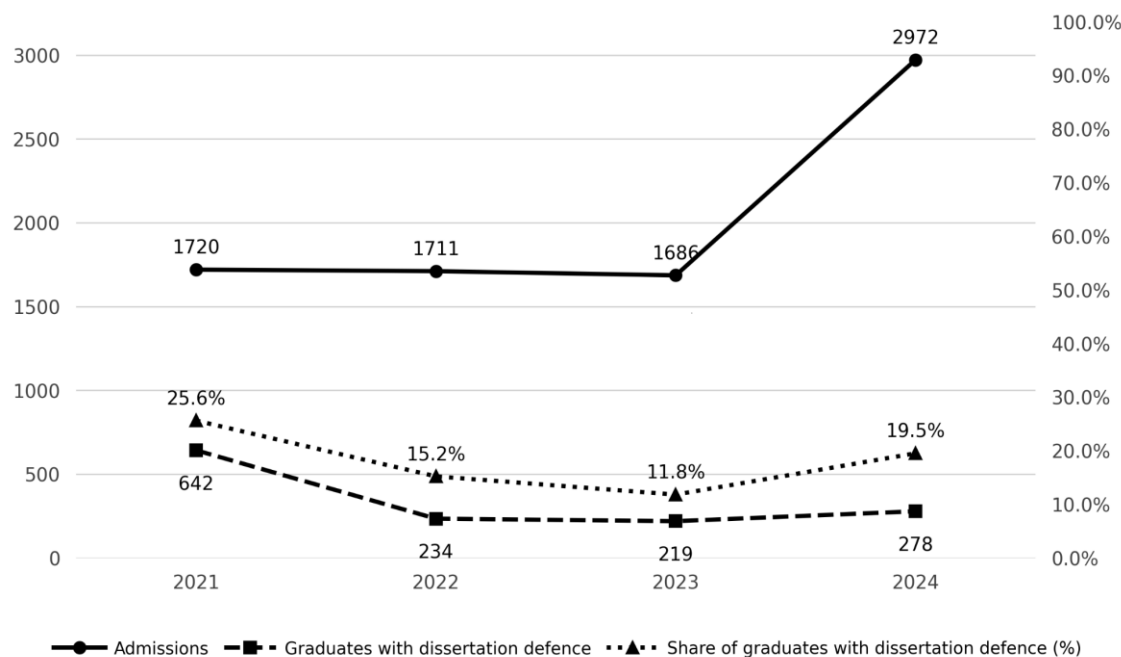
Official statistics indicate substantial but uneven change in doctoral education in Kazakhstan between 2021 and 2024. Doctoral admissions remained relatively stable between 2021 and 2023, at 1,720, 1,711 and 1,686 respectively, before increasing sharply to 2,972 in 2024. Overall, admissions were approximately 73% higher in 2024 than in 2021 (National Academy of Sciences of the Republic of Kazakhstan under the President of the Republic of Kazakhstan, 2025).

Patterns of doctoral completion within the normative period differed from those of admissions. The number of doctoral graduates declined from 2,503 in 2021 to 1,536 in 2022, increased to 1,854 in 2023, and then fell to 1,422 in 2024. The number of graduates completing with dissertation defence also declined substantially, from 642 in 2021 to 234 in 2022 and 219 in 2023, before increasing to 278 in 2024.

As a proportion of all doctoral graduates, those graduating with dissertation defence accounted for approximately 25.6% in 2021, 15.2% in 2022, 11.8% in 2023, and 19.5% in 2024. Thus, the proportion partially recovered in 2024 after declining in the preceding two years but remained below its 2021 level. These percentages are calculated from the official graduation data and should not be interpreted as cohort completion rates.

Figure 2.

National Trends in Doctoral Education in Kazakhstan, 2021–2024



Note. Admissions and graduates with dissertation defence are shown as annual counts. The percentage series represents graduates with dissertation defence as a proportion of total doctoral graduates in each year. These indicators are annual system-level measures and should not be interpreted as a longitudinal cohort completion rate.

Source. National Academy of Sciences of the Republic of Kazakhstan (2025), Table 2.13, based on Bureau of National Statistics data.

Taken together, the results show that changes in doctoral participation, graduation, and dissertation defence did not follow a common trajectory between 2021 and 2024. The national statistics cannot explain the causes of these differences, but they indicate that growth in admissions does not necessarily correspond with parallel changes in other system-level indicators. These patterns provide the empirical basis for examining the organisational conditions through which universities support PhD provision during periods of expansion.

Discussion

Supervision and research environments

The national statistics do not provide evidence about the quality of doctoral supervision in Kazakhstan. They do, however, make supervisory capacity an important consideration in the context of expanding doctoral admissions. Larger doctoral cohorts increase demand for experienced supervisors, regular feedback, mentoring, academic guidance, and opportunities for research collaboration. The implications of such growth depend not only on the number of available supervisors but also on how universities organise and support supervisory work through workload allocation, professional development, co-supervision, mentoring, and quality assurance (Halse & Malfroy, 2010; McAlpine, 2013; Guarimata-Salinas et al., 2024).

This perspective shifts attention from individual supervisors towards institutional responsibility. Maintaining consistent supervisory support during expansion requires organisational systems capable of responding to growth in doctoral participation. The sharp increase in admissions observed in 2024 does not demonstrate that existing supervisory capacity is insufficient, but it reinforces the importance of examining whether institutional support develops alongside increases in doctoral numbers.

The same consideration applies to research environments. Doctoral development extends beyond the supervisor–candidate relationship. Participation in research groups, seminars, collaborative projects, conferences, and scholarly networks provides candidates with opportunities for intellectual exchange, methodological development, and integration into research communities (McAlpine & Norton, 2006;

Rönkkönen et al., 2024). The concept of critical mass is therefore particularly relevant, as universities differ in the concentrations of research expertise, activity, and infrastructure available to doctoral candidates (EUA Council for Doctoral Education, 2016).

The national data cannot assess these institutional differences. Nevertheless, they highlight an important distinction: increasing the number of doctoral places is not equivalent to increasing doctoral research capacity. Sustainable expansion requires attention to the supervisory and research environments within which additional doctoral candidates are expected to develop as researchers.

Governance, progress monitoring, and research integrity

Expansion also increases organisational demands associated with programme governance and doctoral progression. National regulation can establish degree requirements and minimum standards, but universities remain responsible for translating these requirements into programme organisation, academic oversight, candidate support, and quality assurance. Structured annual reviews, milestone assessments, doctoral committees, and early intervention mechanisms can strengthen institutional oversight of candidate progression and reduce reliance on informal supervisory arrangements alone (McAlpine, 2013; Skopek et al., 2022).

The national patterns of graduation and dissertation defence do not demonstrate that progress-monitoring arrangements are inadequate in Kazakhstan. However, the proportion of graduates completing with dissertation defence was lower in 2024 than in 2021, despite some recovery from the lower levels recorded in 2022 and 2023. This suggests the value of complementing aggregate national statistics with information on how universities monitor doctoral progression, respond to delays, and coordinate academic support.

Such evidence would allow policy makers to distinguish more clearly between system-level indicators and the institutional processes underlying them. Measures of supervisory workload, milestone completion, doctoral participation in research, and access to academic support could complement existing indicators of admissions and graduation.

Research integrity forms a further part of this institutional environment. Responsible research is influenced not only by individual conduct but also by university policies, publication expectations, ethics education, assessment procedures, supervisory support, and academic cultures (Abdi et al., 2021). This issue is particularly relevant in Kazakhstan, where publication requirements have become an important feature of doctoral education and have been shown to interact with academic incentives, institutional expectations, and research practices (Kuzhabekova, 2025; Bayanbayeva, 2026; Shamatov et al., 2026).

Including research integrity within the framework does not imply that the observed patterns of graduation or dissertation defence reflect integrity problems. Rather, it recognises that high-quality doctoral provision depends on institutional environments capable of supporting responsible research while managing the academic pressures associated with publication and doctoral progression.

From system expansion to institutional development

The central interpretation emerging from the analysis is that the development of PhD provision is neither linear nor uniform. The findings point to an uneven relationship between expansion in doctoral participation and patterns of graduation and dissertation defence. This does not in itself indicate deterioration in the PhD system. Rather, it suggests that different dimensions of system development may move at different rates.

This distinction can be understood in terms of two related but analytically separate processes. The first is system expansion, reflected in regulatory reform, increased opportunities for PhD study, and growth in admissions. The second is institutional development, reflected in universities' capacity to sustain supervision, research environments, governance, progress monitoring, and research integrity.

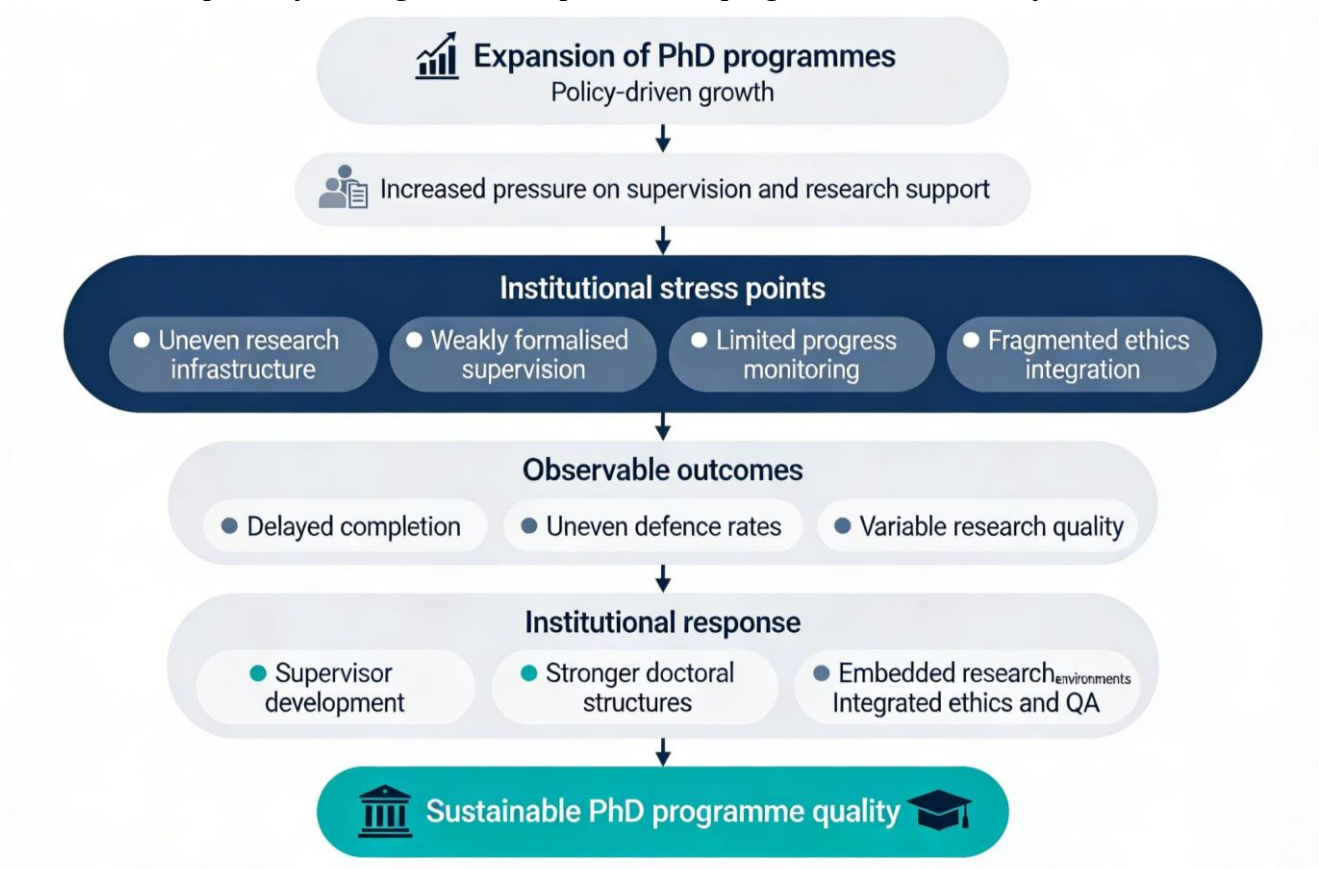
These processes need not develop at the same pace. Changes in regulation and the number of available PhD places can occur relatively rapidly, whereas experienced supervisory communities, mature research environments, effective governance systems, and strong research cultures develop cumulatively over time (Skopek et al., 2022; Taylor, 2023). Evidence from Uzbekistan similarly illustrates how expansion and reform of PhD provision can interact with institutional capacity constraints and organisational practices (Ruziev et al., 2026).

The Kazakhstan data do not demonstrate that institutional capacity has failed to keep pace with expansion, nor do they establish a causal relationship between organisational conditions and PhD outcomes. They show instead that different components of the national PhD system have followed different trajectories during a period of expansion. Institutional capacity therefore becomes an important object of empirical

investigation: it provides a basis for asking how universities translate system-level growth into sustainable conditions for PhD training and research.

Figure 3.

Institutional pathway linking doctoral expansion and programme sustainability in Kazakhstan



Implications for policy and research

The analysis has three main implications. First, national monitoring of doctoral education could complement admissions and graduation statistics with indicators of institutional capacity. Relevant measures could include supervisory resources and workloads, doctoral participation in research environments, progress-monitoring arrangements, and access to academic and research support. Such information would help distinguish growth in doctoral provision from development of the organisational capacity required to sustain it.

Second, further expansion in doctoral admissions should be considered alongside institutional readiness. This does not imply restricting doctoral growth. Rather, expansion should be accompanied by attention to the supervisory resources, research environments, governance arrangements, and academic support available within universities.

Third, the institutional capacity framework proposed in this study requires empirical testing. Future research should examine variation across universities and disciplines using institutional administrative data, interviews with doctoral candidates and supervisors, analysis of programme governance, and comparative case studies. Longitudinal research would be particularly valuable for examining how institutional capacity develops in relation to changes in doctoral participation, progression, graduation, and dissertation defence.

Conclusion

Kazakhstan has substantially expanded PhD provision while continuing to reform the regulatory framework governing PhD programmes. The national trends examined in this study show that different dimensions of the system have not followed the same trajectory: admissions remained relatively stable between 2021 and 2023 before increasing sharply in 2024, while graduation and dissertation defence followed different patterns. These findings reinforce the need to examine PhD development not only through system-level

expansion but also through the organisational conditions that support it.

This article proposes institutional capacity as a framework for such analysis. It brings together five interrelated organisational domains as supervision, research environments, governance, progress monitoring, and research integrity, and distinguishes system expansion from the institutional development required to sustain it. The framework is interpretive rather than directly measured, but it identifies organisational conditions that aggregate national statistics cannot capture.

The practical implication is that regulatory reform and growth in doctoral participation should be accompanied by attention to universities' supervisory capacity, research environments, academic support, governance, and research integrity systems. Future research should test these relationships using institution-level, comparative, and longitudinal evidence. Although developed through the Kazakhstan case, the distinction between system expansion and institutional development may also provide a useful perspective for analysing doctoral education in other higher education systems experiencing rapid growth.

Conflict of Interest Statement

The author declare no potential conflicts of interest regarding the research, authorship, or publication of this article.

Declaration of the use of AI

The author declare that no artificial intelligence (AI) tools were used in the preparation of this manuscript.

References

- Abdi, S., Pizzolato, D., Nemery, B., & Dierickx, K. (2021). Educating PhD students in research integrity in Europe. *Science and Engineering Ethics*, 27, Article 5. <https://doi.org/10.1007/s11948-021-00290-0>
- Bayanbayeva, A. (2026). The impact of the 'publish or perish' culture on research practices and academic life in Kazakhstan: Challenges and consequences in the age of global university rankings. *Higher Education Research & Development*, 45(1), 1–15. <https://doi.org/10.1080/07294360.2025.2526051>
- EUA Council for Doctoral Education. (2016). Doctoral education: Taking Salzburg forward – Implementation and new challenges. European University Association.
- Guarimata-Salinas, G., Carvajal, J. J., & Jiménez López, M. D. (2024). Redefining the role of doctoral supervisors: A multicultural examination of labels and functions in contemporary doctoral education. *Higher Education*, 88, 1305–1330. <https://doi.org/10.1007/s10734-023-01171-0>
- Halse, C., & Malfroy, J. (2010). Retheorizing doctoral supervision as professional work. *Studies in Higher Education*, 35(1), 79–92. <https://doi.org/10.1080/03075070902906798>
- Ibraev, A. Zh., Kul'evskaya, Y. G., Ulez'ko, G. G., & Galants, E. A. (2015). Reforming the system of research-staff training: Doctoral (Ph.D.) education in Kazakhstan. *Scientific and Technical Information Processing*, 42, 78–84. <https://doi.org/10.3103/S0147688215020082>
- Kuzhabekova, A. (2020). Development and transformation of doctoral education in Kazakhstan. In M. Yudkevich, P. G. Altbach, & H. de Wit (Eds.), *Trends and issues in doctoral education: A global perspective* (pp. 340–362). SAGE Publications. <https://doi.org/10.4135/9789353885991.n14>
- Kuzhabekova, A. (2025). Ph.D. publication requirement and its effects on research productivity trends in Kazakhstan. *Higher Education Quarterly*, 79(1), Article e12590. <https://doi.org/10.1111/hequ.12590>
- Lee, A. (2008). How are doctoral students supervised? Concepts of doctoral research supervision. *Studies in Higher Education*, 33(3), 267–281. <https://doi.org/10.1080/03075070802049202>
- McAlpine, L. (2013). Doctoral supervision: Not an individual but a collective institutional responsibility. *Journal for the Study of Education and Development*, 36(3), 259–280. <https://doi.org/10.1174/021037013807533061>
- McAlpine, L., & Norton, J. (2006). Reframing our approach to doctoral programs: An integrative framework for action and research. *Higher Education Research & Development*, 25(1), 3–17. <https://doi.org/10.1080/07294360500453012>
- National Academy of Sciences of the Republic of Kazakhstan under the President of the Republic of Kazakhstan. (2025). Report of the National Academy of Sciences of the Republic of Kazakhstan under

- the President of the Republic of Kazakhstan on Science for 2024. Almaty. ISBN 9965-25-129-0.
- Rönkkönen, S., Tikkanen, L., Virtanen, V., & Pyhältö, K. (2024). The impact of supervisor and research community support on PhD candidates' research engagement. *European Journal of Higher Education*, 14(4), 536–553. <https://doi.org/10.1080/21568235.2023.2229565>
- Ruziev, K., Rabbimova, D., & Mamasolieva, M. (2026). Reforming doctoral education in Uzbekistan: Models, rules and norms of behaviour. *Higher Education Policy*, 39, 500–526. <https://doi.org/10.1057/s41307-025-00396-1>
- Satpayeva, Z., Kangalakova, D., Kakizhanova, T., & Konstantinova, K. (2025). Challenges of Ph.D. training in social sciences and humanities in Kazakhstani universities. *Knowledge and Performance Management*, 9(1), 1–15. [https://doi.org/10.21511/kpm.09\(1\).2025.01](https://doi.org/10.21511/kpm.09(1).2025.01)
- Shamatov, D., Ispambetova, B., Baikenov, K., & Kuzhabekova, A. (2026). Research integrity under publication requirements: A qualitative study of doctoral education in Kazakhstan. *International Journal for Educational Integrity*, 22, Article 9. <https://doi.org/10.1007/s40979-026-00219-9>
- Skopek, J., Triventi, M., & Blossfeld, H.-P. (2022). How do institutional factors shape PhD completion rates? An analysis of long-term changes in a European doctoral program. *Studies in Higher Education*, 47(2), 318–337. <https://doi.org/10.1080/03075079.2020.1744125>
- Taylor, S. (2023). The changing landscape of doctoral education: A framework for analysis and introduction to the special issue. *Innovations in Education and Teaching International*, 60(5), 606–622. <https://doi.org/10.1080/14703297.2023.2237962>
- Tokatligil, Y., Saliyeva, A., Mambetalina, A., Nurmanbetova, J., & Karmelyuk, A. (2025). Co-supervision in Kazakhstan PhD programs: The Kazakhstani + international co-supervisor model. *International Journal of Doctoral Studies*, 20, Article 15. <https://doi.org/10.28945/5626>

Information about authors

Aida Sagintayeva – PhD, Vice Provost for Academic Planning and Budget, Dean of the Graduate School of Education, Nazarbayev University, Astana, Republic of Kazakhstan, asagintayeva@nu.edu.kz, ORCID iD: 0000-0001-6249-0840 (*corresponding author*)

Gulmira Qanay – PhD, Graduate School of Education, Nazarbayev University, Astana, Republic of Kazakhstan, gkanayeva@nu.edu.kz, ORCID: 0000-0003-4935-6027

Received: 21.03.2026

Revised: 29.05.2026

Revised: 12.08.2026

Accepted: 25.08.2026

Published: 30.06.2026

Type: Research Article

Zhibek Tleshova*, Nurgissa Kabarbek, Bexultan Turgunov

Astana IT University, Astana, Kazakhstan

THE ELECTRONIC SYLLABUS AS A PEDAGOGICAL OBJECT: AN INTEGRATIVE LITERATURE REVIEW

Abstract. Every course delivered through a learning management system (LMS) opens with an electronic syllabus. Although what we mean by electronic syllabus is rarely examined on its own terms. This integrative literature review attempts to group research papers on electronic syllabus that have developed largely in isolation: 1) research on the syllabus as a pedagogical document, 2) research on learning management systems and technology integration, and 3) the literature on the digital transformation of Kazakhstani higher education. The synthesis of the reviewed literature groups the evidence into five pedagogical clusters: alignment of outcomes, activities and assessment; transparency of expectations; learner-centred framing and tone; structure, navigation and multimodality; and governance, standardisation and analytics. From these clusters the article proposes a three-layer model of the electronic syllabus, comprising design, communication and use. Astana IT University serves as an illustrative setting but not a site of field research. This article concludes with a research agenda built around document analysis of LMS-hosted syllabi and studies of their use.

Keywords: electronic syllabus; integrative literature review; learning management system; constructive alignment; learner-centred pedagogy; higher education; Kazakhstan

Introduction

Over the past two decades the learning management system has become the organising infrastructure of university teaching: enrolment, readings, announcements, submissions and grades pass through a single platform, which now forms an integral part of the learning environment in most institutions (Turnbull et al., 2022). The syllabus inside that environment in its digital form is usually the first structured document a student sees. In practice the electronic syllabus is handled as a template to be completed rather than as an act of teaching.

A well-written syllabus makes visible the constructive alignment between the intended learning outcomes, activities, and the assessment that tests the attained outcomes (Biggs, 1996; Grunert O'Brien et al., 2008), and the instructor's tone in the syllabus shapes students' perception about the course and the instructor before teaching has begun (Harnish & Bridges, 2011). Placing a digital syllabus in an LMS is expected to strengthen these functions. Inside the platform the syllabus can link directly to the readings and assessments, and connects to the assessment policy and to the course content. In practice, however, it is usually held in a single long template through which students scroll, so the connections between outcomes, activities and assessment are presented one section at a time rather than seen together. Whether a template of this kind preserves the pedagogical function of the syllabus remains unexamined.

The electronic syllabus, though produced inside this digital infrastructure, has yet to become an object of study at institutional level in Kazakhstan. Dzhanegizova (2024) reports that platform coverage reached every Kazakhstani university by 2023, up from a quarter in 2015, but that the capacity to use those platforms did not spread evenly: national universities scored 4.2 out of 5 on digital readiness, regional institutions 3.5 and private ones 3.1. Astana IT University (AITU) illustrates this pattern. Founded in 2019 as an autonomous, English-medium institution charged with advancing digital transformation in Kazakhstan, it delivers much of its teaching in blended and technology-mediated forms (Mazhit et al., 2023; Zhanarstanova et al., 2024) and builds its courses inside the platform, so that its electronic syllabi are part of the study environment rather than attached to it.

This study attempts to do an integrative literature review in order to summarize research papers to achieve a comprehensive understanding of what electronic syllabus meant to be in digital infrastructure at higher education institution level. This study helps put together published papers relevant to the electronic

syllabus, organises it into clusters, and uses that organisation to define what the pedagogical function of an electronic syllabus involves, to identify where evidence is absent, and to specify the studies that needs further research. Three questions are set to flow our study:

RQ1. How does the existing literature characterise the pedagogical functions of the course syllabus?

RQ2. How are those functions hosted inside a learning management system as a part of digital infrastructure?

RQ3. What gaps in the literature exist and what features of the Kazakhstani setting define a priority for future research?

Literature Review

The syllabus as a pedagogical document. The meaning of electronic syllabus (further e-syllabus) is first to be understood and defined from the outset of this study. A Cambridge online dictionary defines “a syllabus as a plan showing the subjects or books to be studied in a particular course” (Cambridge Online Dictionary). That covers what the document lists, but not what it does. Parkes and Harris (2002) distinguish three purposes that pull in different directions. They are the syllabus as a record used for accreditation and accountability, as a contract setting out mutual expectations, and as a cognitive map helping students grasp how a course is set up. A recent scoping review finds that the syllabus still serves these competing uses, and that the tensions between them remain unresolved (Gauthier et al., 2026).

An e-syllabus is the document when it is built inside a learning management system. A syllabus written elsewhere and uploaded as a file is a printed syllabus behind a login. Izienicki (2026) rebuilt a course syllabus as a set of LMS modules rather than as a single linear file, and found that students consulted the modular version twice as often across a semester.

Putting together these perspectives suggest that a syllabus is more than a course plan or an administrative document. It is a pedagogical document that communicates the goals, structure, expectations, and learning route of a course. It also shapes students’ understanding of how learning will take place. The scholarly literature views the syllabus as a tool for supporting student engagement and success as well. Within this study, therefore, an e-syllabus is defined as a syllabus that is created and delivered within a learning management system, where its elements can be organized, linked, updated, and accessed interactively. An e-syllabus functions as an integrated digital environment that supports navigation, communication, and student engagement throughout the course.

In the modern educational settings, the syllabus is expected to have a function as a learning tool rather than a note pad. The syllabus belongs to the classroom, where it is employed by teachers to work with learners, and the institution, which has requirements of its own (Gauthier, G., Banner, J., & Winer, L., 2026) Grunert O'Brien et al. (2008) argued that a syllabus should explain not only a course's structure but its teaching philosophy and the reasoning behind its activities. Weimer (2013) describes learner-centred teaching as a series of shifts in who holds power, what content is for, and the purpose of assessment, each of which a syllabus either displays or conceals, and Fink (2013) argues that a course, and by extension its syllabus, should be planned backwards from significant learning goals rather than forwards from a list of topics (see also Bain, 2004).

Parkes and Harris (2002) provide a useful framework for understanding the multiple functions of the syllabus. They describe it as a record, serving accreditation and accountability purposes; a contract, outlining the responsibilities and expectations of both instructors and students; and a cognitive map, helping students understand the structure, goals, and coherence of a course. While all three functions remain important, they can compete with one another. The record function prioritizes completeness and compliance with institutional requirements, whereas the contract function often encourages detailed policies and regulations. By contrast, the cognitive-map function emphasizes the syllabus as a pedagogical tool that supports student learning. As institutional policies on attendance, assessment, academic integrity, and course procedures accumulate over time, syllabi can become increasingly administrative and less focused on learning. In digital learning environments, these priorities are often embedded in standard LMS templates. Consequently, the design of the syllabus, including its digital form, should be viewed not merely as an administrative task but as a pedagogical decision that influences how students engage with and navigate a course.

This pedagogical perspective has also led researchers to develop tools for evaluating the quality of

syllabi. Slattery and Carlson (2005) emphasize that a syllabus should do more than provide information; it should also engage and motivate students. Building on this idea, Cullen and Harris (2009) developed a rubric for assessing how learner-centred a syllabus is, examining such aspects as the distribution of responsibility between instructor and students, the presentation of course content, and the role of assessment. Later, Palmer et al. (2014) expanded this work by proposing a learning-focused rubric that places greater emphasis on supporting student learning. These studies suggest that the pedagogical value of a syllabus can be systematically evaluated rather than assumed. However, despite the growing research on learner-centred syllabi, similar evaluation frameworks have not yet been applied to syllabi designed and delivered within LMS environments.

Research has also shown that the way a syllabus is written can influence how students experience a course. For example, Harnish and Bridges (2011) found that when the wording of a syllabus was more welcoming and supportive, students perceived the instructor as more approachable, even though the course content remained unchanged. This finding suggests that the syllabus communicates not only information but also the tone and culture of a course. Similarly, research of transparency of syllabi has demonstrated the importance of clearly explaining the purpose of learning activities, what students are expected to do, and how their work will be evaluated. Such clarity has been associated with improved academic performance and has been particularly beneficial for students who may be less familiar with university expectations. Although this work has often focused on assignments and assessment tasks, its principles are equally relevant to the syllabus, which is usually students' first guide to a course.

The syllabus is also the place where the alignment between course aims, learning activities, and assessment becomes visible to students. In this sense, it gives practical expression to the principle of constructive alignment (Biggs, 1996). Recognising this role, Nilson (2007) proposed the use of graphic syllabi and outcomes maps to help students see how different elements of a course are connected. While originally designed for printed documents, these visual approaches anticipate many of the possibilities offered by contemporary learning management systems. The syllabus is not a fixed or purely administrative document. Rather, it continues to evolve in response to new educational challenges and increasingly serves as a space where expectations, responsibilities, and course policies are negotiated and communicated.

The second part of the literature review concerns the platform. Researchers describe the LMS as a platform that integrates course content, communication, assessment, and learning activities (Akram et al., 2021). At the same time, the literature suggests that the effectiveness of an LMS depends less on its technical features than on how it is used by instructors and supported by institutions. Studies conducted in rapidly digitalizing higher education systems indicate that factors such as perceived usefulness, institutional support, and staff training are stronger predictors of successful LMS implementation than technological capability alone (Sulaiman, 2024). As a result, many institutions continue to use their LMS primarily as a repository for course documents rather than as an interactive learning environment.

Explanations of LMS adoption are often based on technology acceptance models. The Technology Acceptance Model (Davis, 1989) highlights perceived usefulness and ease of use, while the Unified Theory of Acceptance and Use of Technology adds factors such as performance expectations, social influence, and institutional support (Venkatesh et al., 2003). Applied to the e-syllabus, these models help explain whether students and instructors consider it useful enough to use regularly. However, they say little about its educational value. This gap is addressed by the Technological Pedagogical Content Knowledge (TPACK) framework, which argues that effective technology use depends on the interaction of technological, pedagogical, and subject knowledge (Mishra & Koehler, 2006). The framework continues to inform research on technology integration in higher education (Kholid et al., 2023; Gonscherowski & Rott, 2025).

The third group of the literature focuses on the Kazakhstani context. Higher education digitalisation in Kazakhstan has progressed rapidly, driven largely by national initiatives. Studies report positive links between digitalisation and educational quality, but also point to continuing challenges, particularly limited digital competence among faculty and resistance to organisational change (Dzhanegizova, 2024). Recent research views digital transformation as more than the adoption of technology, emphasizing changes in institutional structures, strategies, teaching practices, and culture (Zhukabayeva et al., 2025). At the same time, reviews of national digitalisation projects show that research on their pedagogical impact remains relatively limited (Narbaev et al., 2025).

At AITU, existing studies report generally positive outcomes from blended learning and have begun to use platform data to examine teaching and learning processes (Mazhit et al., 2023; Zhanarstanova et al., 2024). However, interest in learning analytics is still emerging, and little attention has been paid to how courses are actually presented to students through the LMS.

Taken together, these three groups of research point to a common gap. Studies on syllabi emphasize their pedagogical importance, while research on LMSs focuses mainly on technology adoption and use. In Kazakhstan, attention has largely been directed towards digital platforms and educational digitalisation. Yet the e-syllabus remains largely unexplored in all three areas. Although it is a standard feature of virtually every course, little research has examined how it is designed, used, or experienced by students and instructors.

Conceptual Framework

To understand the e-syllabus as a teaching tool rather than an administrative document, this study draws on three perspectives. Constructive alignment explains how different elements of a course should fit together, but it does not explain whether students actually use the syllabus. Technology acceptance models focus on use and adoption, but pay little attention to pedagogy. The transparency framework focuses on clarity and student understanding, but does not address either alignment or use.

From the perspective of constructive alignment, a syllabus should clearly connect learning outcomes, learning activities, and assessment (Biggs, 1996). The transparency framework highlights the importance of explaining the purpose of the course, student responsibilities, and assessment criteria in language that is clear and accessible (Winkelmes et al., 2016). Research also shows that a welcoming tone can positively influence students' perceptions of a course and its instructor (Harnish & Bridges, 2011). Finally, technology acceptance models and the TPACK framework help explain whether instructors and students find the e-syllabus useful enough to engage with regularly (Davis, 1989; Venkatesh et al., 2003; Mishra & Koehler, 2006).

Based on these perspectives, the e-syllabus can be understood through three interconnected dimensions: design, communication, and use. The design dimension concerns how course outcomes, activities, and assessments are organised and connected. The communication dimension focuses on the clarity, transparency, and tone of the document. The use dimension concerns whether students and instructors actually access and engage with it. The LMS provides the environment in which all three dimensions operate, offering opportunities for linking resources, updating information, and tracking engagement.

These dimensions are closely related. A syllabus that lacks clear alignment cannot be improved simply through friendly language, while a well-designed syllabus has little educational value if students do not use it. At the same time, patterns of use can demonstrate strengths and weaknesses in design and communication. For this reason, the pedagogical quality of an e-syllabus should be considered as the interaction of design, communication, and use rather than as separate elements. This framework guides the analysis presented in the following sections.

Table 1.

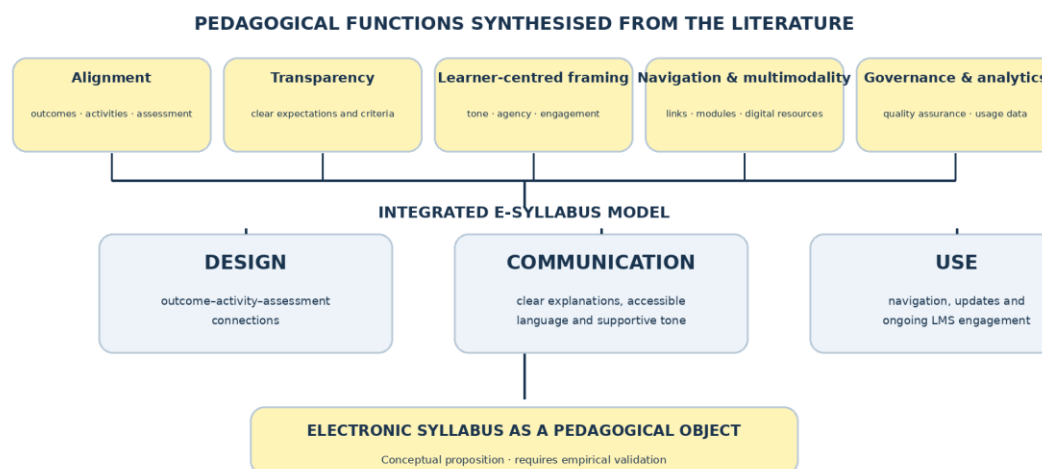
Conceptual dimensions of the electronic syllabus

Dimension	Focus identified in the literature	LMS-related manifestation
Design	Constructive alignment of learning outcomes, activities and assessment	Links among outcomes, course materials, assignments and rubrics
Communication	Transparency of expectations and learner-centred tone	Accessible policies, clear language, guidance and regular updates
Use	Access and meaningful engagement throughout the course	Navigation, interactivity and evidence of student/instructor use

Source: authors' conceptual synthesis of the literature reviewed in this article.

Figure 1.

Conceptual synthesis of the electronic syllabus as a pedagogical object.



Note. The five thematic clusters are organised under three interrelated dimensions. The proposed relationships are interpretive and have not been empirically tested.

Methodology

Research design

This study adopts an integrative literature review to bring together theoretical and empirical scholarship from three related fields: the syllabus as a pedagogical document, learning management systems and technology integration, and digital transformation in Kazakhstani higher education. An integrative approach is appropriate because the literature addresses these dimensions through different disciplinary and methodological perspectives. The review seeks to develop a conceptual understanding of the electronic syllabus as a pedagogical object rather than estimate an intervention effect or claim exhaustive coverage of all published studies (Torraco, 2005; Grant & Booth, 2009; Snyder, 2019). Astana IT University is considered as an illustrative institutional context on the basis of published and institutional sources; no primary research involving AITU students, staff or course documents was conducted.

Literature search and selection

As documented in the source manuscript, literature was identified through Scopus, Web of Science, ERIC and Google Scholar, complemented by targeted searches of regional sources and relevant reference lists. Search concepts combined terms relating to the syllabus or electronic syllabus, learning management systems and pedagogy. The material considered included peer-reviewed publications, scholarly books, and institutional and policy documents relevant to syllabi, LMS use in higher education or digital transformation in Kazakhstan. Foundational theoretical contributions were considered alongside empirical research, with priority given to recent publications. Materials addressing only technical aspects of platforms, non-scholarly sources and texts unavailable in full were excluded. These procedures reflect the selection approach described by the authors and should not be interpreted as a fully documented systematic search.

Data extraction and thematic synthesis

The selected publications were examined with reference to their research focus, theoretical or methodological approach, principal findings and relevance to the pedagogical functions of the syllabus. The review applies a thematic-synthesis approach informed by Braun and Clarke (2006): related findings were compared across the three literature streams, recurring themes were identified and thematically related findings were brought together. This synthesis is presented through five pedagogical clusters: (1) alignment of learning outcomes, activities and assessment; (2) transparency of expectations; (3) learner-centred framing and tone; (4) structure, navigation and multimodality; and (5) governance, standardisation and analytics. The clusters provide the basis for interpreting the three interconnected dimensions of the proposed conceptual model:

design, communication and use. This is an interpretive synthesis of the cited scholarship, not a statistical aggregation of study findings.

Table 2.
Selected literature underpinning the integrative synthesis

Study/source	Focus or approach	Contribution to this review
Parkes & Harris (2002)	Conceptual discussion of syllabus purposes	Record, contract and cognitive-map functions; foundational framing of the syllabus as more than a course outline.
Biggs (1996)	Constructive alignment in higher education	Links intended learning outcomes, learning activities and assessment; informs the design dimension.
Cullen & Harris (2009); Palmer et al. (2014)	Rubric-based examination of course syllabi	Offers criteria for learner-centred and learning-focused syllabus design.
Harnish & Bridges (2011)	Study of syllabus tone and student perceptions	Supports the relevance of welcoming language for perceived instructor approachability.
Winkelmes et al. (2016)	Teaching intervention concerning transparency	Provides a basis for examining clarity of purposes, tasks and assessment expectations.
Davis (1989); Venkatesh et al. (2003)	Technology acceptance models	Highlights perceived usefulness and adoption, relevant to the use dimension.
Mishra & Koehler (2006)	TPACK conceptual framework	Connects technology use with pedagogical and content knowledge.
Turnbull et al. (2022); Akram et al. (2021)	LMS policies and technology integration	Provides context for the digital environment in which an electronic syllabus is delivered.
Eslami et al. (2024)	Empirical study of syllabus learner-centredness	Links syllabus characteristics to differences in student outcomes in STEM higher education.
Dzhanegizova (2024); Mazhit et al. (2023); Zhanarstanova et al. (2024)	Kazakhstan digital transformation, blended learning and analytics	Supplies national and AITU-related context; does not constitute direct evidence on AITU e-syllabi.
Gauthier et al. (2025/2026)	Scoping review of syllabus uses	Identifies tensions between administrative, contractual and pedagogical roles; bibliographic year/version requires reconciliation.

Source: authors' synthesis of the studies cited in the manuscript; entries are selected illustrative sources, not a complete inventory of screened records.

Findings

The analysis grouped the literature into five clusters. That the syllabus sustains several competing uses at once is confirmed by a recent scoping review, which finds the document pulled between administrative, contractual and pedagogical functions across the literature as a whole (Gauthier et al., 2025). The first two belong to the design and communication dimensions; the third develops the communication dimension; the fourth describes what the platform adds; the fifth a dimension largely new to the digital form. Together they answer RQ1 and RQ2.

The literature consistently shows that the syllabus is more than an administrative document. Its primary pedagogical role is to make the structure of a course visible by linking learning outcomes, learning activities, and assessment (Biggs, 1996). In an LMS, these relationships can be made explicit through links between outcomes, assignments, rubrics, and gradebook elements. However, this potential depends on design. Studies using syllabus evaluation rubrics have found that many syllabi still lack clear connections between outcomes, learning activities, and assessment (Cullen & Harris, 2009; Palmer et al., 2014; Serbati et al., 2022). The literature therefore suggests that digital delivery alone does not improve alignment; effective use requires the technological and pedagogical capabilities described by TPACK (Mishra & Koehler, 2006). Research on

electronic syllabi in Kazakhstan remains absent.

A second theme concerns transparency. Making course purposes, tasks, and assessment criteria explicit improves student learning and particularly benefits those who are less familiar with academic expectations (Winkelmes et al., 2016). LMS platforms can support transparency by placing relevant information alongside learning activities and making it available throughout the semester. At the same time, digital environments can create new challenges when information is dispersed across multiple pages or modules. Transparency therefore depends not only on what information is provided but also on how it is organised. This issue is particularly relevant in multilingual educational settings, where the clarity of course information can directly affect students' access to learning.

The literature also highlights the importance of learner-centred approach. Rather than serving only as a record or contract, the syllabus can shape students' perceptions of the course and instructor. More supportive and learner-centred language has been associated with greater approachability and more positive learning experiences (Harnish & Bridges, 2011). Recent evidence further suggests that learner-centred syllabi may contribute to reducing achievement gaps among different student groups (Eslami et al., 2024). Within an LMS, however, institutional templates can create tension between consistency and personalisation. While standardisation of syllabi supports quality assurance, it can also limit the instructor's voice and reduce opportunities for meaningful engagement (Judson, 2025).

The electronic syllabus differs most clearly from its printed predecessor through its capacity for navigation, interactivity, and multimodal presentation. Unlike a static document, it can incorporate hyperlinks, embedded resources, visual elements, and continual updates. These features can strengthen alignment, transparency, and learner engagement, while also supporting broader course design principles associated with effective online learning (Nilson, 2007; Khodabandelou et al., 2025). As a result, the pedagogical potential of the electronic syllabus often remains not complete.

Finally, the literature points to a growing governance function of the electronic syllabus. LMS-based syllabi can be standardised, monitored, and incorporated into quality assurance processes. They can also generate data about student engagement and access, an area that has begun to attract attention in Kazakhstan (Dzhanegizova, 2024; Zhanarstanova et al., 2024). While standardisation can improve consistency across courses, excessive emphasis on compliance may weaken the syllabus's pedagogical functions. In the terms of Parkes and Harris (2002), digital systems often strengthen the syllabus as a record and a contract, while giving less attention to its role as a learning tool. This tension is becoming more important as institutions increasingly use syllabi to communicate policies on emerging issues such as generative artificial intelligence (An et al., 2025).

Taken together, these findings suggest that the pedagogical development of an electronic syllabus involves more than transferring a traditional document into an LMS. An effective e-syllabus aligns outcomes, activities, and assessment; communicates expectations clearly; supports learner engagement; uses the interactive features of the platform; and balances institutional requirements with educational goals. Across all three research questions, one conclusion emerges consistently: although the e-syllabus has become a routine component of digital higher education, it remains underexplored.

Table 3.

Summary of the five pedagogical clusters emerging from the integrative synthesis

Cluster	Interpretation based on the reviewed literature
Alignment	Explicit connections among outcomes, activities, assessment and rubrics.
Transparency	Accessible expectations, assessment criteria and course policies.
Learner-centred framing	Supportive tone, student agency and approaches to the instructor.
Navigation and multimodality	LMS links, modules, visuals, embedded resources and accessible updates.
Governance and analytics	Standardised templates, quality assurance, access patterns and engagement data.

Source: synthesis of the five clusters described in the Findings section.

Discussion

This review addressed three questions concerning the electronic syllabus as a pedagogical object. First, the literature consistently characterises the syllabus as more than an administrative document. Its key pedagogical functions are to communicate course structure, clarify expectations, support student engagement, and connect learning outcomes, activities, and assessment (Biggs, 1996; Winkelmes et al., 2016; Weimer, 2013). These functions were identified well before the widespread use of learning management systems and remain central to syllabus design today.

Second, the findings show that an LMS can either strengthen or weaken these functions. When designed intentionally, electronic syllabi can make course structures more visible, provide easier access to information, support continuous communication, and connect students directly to learning activities and assessment. However, the literature suggests that these benefits do not result from technology alone. Their effectiveness depends on pedagogical design, instructor capability, and institutional support (Mishra & Koehler, 2006; Davis, 1989; Venkatesh et al., 2003). In many cases, LMS platforms continue to be used primarily as repositories for documents rather than as interactive learning environments (Akram et al., 2021).

The main contribution of this review is a shift in perspective. Most institutional approaches evaluate syllabi in terms of completeness, compliance, and consistency. The literature reviewed here suggests that the electronic syllabus should instead be understood as a teaching tool. From this perspective, its quality depends not only on the information it contains but also on how effectively it supports learning through alignment, transparency, learner-centred communication, and navigation. This reframing helps explain why the e-syllabus has received limited attention in research despite being present in virtually every course.

The findings also have implications for policy and practice. For instructional designers and quality assurance units, LMS templates should support pedagogical goals rather than simply ensure compliance. Clear connections between learning outcomes, activities, and assessment, together with clear assessment criteria and learner-friendly communication, should be built into course templates from the outset. The literature from Kazakhstan similarly suggests that successful digitalisation depends not only on technological infrastructure but also on educators' ability to use technology for pedagogical purposes (Dzhanegizova, 2024; Abiltayeva et al., 2025; Abdigapbarova et al., 2025).

Finally, the review identifies a significant research gap. Although digitalisation has become a major priority in higher education, the electronic syllabus remains largely absent from the literature. This gap is particularly evident in Kazakhstan, where research has focused primarily on platform adoption, online learning, and learning analytics (Narbaev et al., 2025; Zhanarstanova et al., 2024). Future studies should therefore examine LMS-hosted syllabi directly, exploring their design, use, and impact on students' learning experiences. For institutions such as AITU, the electronic syllabus offers a practical way to investigate how pedagogical quality is realised within digital learning environments.

Conclusion

Although the syllabus is a standard feature of higher education, the electronic syllabus remains surprisingly underexplored in research. This review has shown that the syllabus serves important pedagogical functions: it helps students understand how a course is organized, clarifies expectations, and shapes the learning relationship between instructor and student. When hosted within a learning management system, these functions can be strengthened through better connections between outcomes, activities, assessment, and learning resources. At the same time, digital delivery introduces new challenges, including fragmented information, excessive standardization, and a greater emphasis on compliance than on learning.

The review also suggests that the value of an electronic syllabus depends less on the technology itself than on how it is designed and used. A well-designed e-syllabus can support alignment, transparency, student engagement, and navigation throughout the course. For this reason, the electronic syllabus should be viewed as a pedagogical tool rather than simply an administrative requirement.

The challenge is no longer whether institutions use digital platforms, but how effectively those platforms support learning. In this context, the electronic syllabus offers a useful lens through which to examine the pedagogical quality of digital education. As a digitally oriented institution, Astana IT University provides a particularly relevant setting for future research on how e-syllabi are designed, used, and experienced by students and instructors. Such research would help address a significant gap in both the regional and

international literature.

Finding Information

This research paper is carried out within the Astana IT University research funding 2025 scholarship. The research title: “An Integrated System for Correlating Syllabi with Library Resources to Optimise Book Availability”.

Conflict of Interest Statement

The authors declare no potential conflicts of interest regarding the research, authorship, or publication of this article.

Author Contributions

All authors contributed equally to all stages of the research, including conceptualization, methodology, data analysis, manuscript preparation, review, and editing. All authors have read and approved the final version of the manuscript.

Declaration of the Use of AI

Generative AI tools were used to assist with the formatting of citations and reference list.

References

- Abdigapbarova, U., Sadirbekova, D., Nishanbayeva, S., & Zhiyenbayeva, N. (2025). The impact of digital hybrid education model on teachers' engagement and academic performance in the context of Kazakhstan. *Scientific Reports*, 15, 17865. <https://doi.org/10.1038/s41598-025-02875-2>
- Abiltayeva, A., Zhumagulova, K., Satayev, M., Yechshzhanov, T., Maimatayeva, A., & Balta, N. (2025). Teacher competencies and digital integration into teaching practices: Perceptions from pre-service biology teachers in Kazakhstan. *Frontiers in Education*, 10, 1628034. <https://doi.org/10.3389/educ.2025.1628034>
- Akram, H., Yingxiu, Y., Al-Adwan, A. S., & Alkhalifah, A. (2021). Technology integration in higher education during COVID-19: An assessment of online teaching competencies through the technological pedagogical content knowledge model. *Frontiers in Psychology*, 12, 736522. <https://doi.org/10.3389/fpsyg.2021.736522>
- An, Y., Yu, J. H., & James, S. (2025). Investigating the higher education institutions' guidelines and policies regarding the use of generative AI in teaching, learning, research, and administration. *International Journal of Educational Technology in Higher Education*, 22(1), 10. <https://doi.org/10.1186/s41239-025-00507-3>
- Bain, K. (2004). *What the best college teachers do*. Harvard University Press.
- Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32(3), 347–364. <https://doi.org/10.1007/BF00138871>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp0630a>
- Cullen, R., & Harris, M. (2009). Assessing learner-centredness through course syllabi. *Assessment & Evaluation in Higher Education*, 34(1), 115–125. <https://doi.org/10.1080/02602930801956018>
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
- Dzhanegizova, A. (2024). Digital transformation of higher education in Kazakhstan: Challenges and solutions. *Economic Annals-XXI*, 209(5–6), 42–55. <https://doi.org/10.21003/ea.V209-05>
- Eslami, M., Denaro, K., Collins, P., Sumarsono, J. M., Dennin, M., & Sato, B. K. (2024). How syllabi relate to outcomes in higher education: A study of syllabi learner-centeredness and grade inequities in STEM. *PLoS ONE*, 19(4), e0301331. <https://doi.org/10.1371/journal.pone.0301331>
- Fink, L. D. (2013). *Creating significant learning experiences: An integrated approach to designing college courses* (2nd ed.). Jossey-Bass.
- Gauthier, G., Banner, J., & Winer, L. R. (2025). What is the syllabus for? Revealing tensions through a scoping review of syllabus uses. *International Journal for Academic Development*, 1–20.

<https://doi.org/10.1080/1360144X.2025.2465363>

- Gonscherowski, P., & Rott, B. (2025). A systematic review of the literature on TPACK instruments used with pre-service teachers from 2017 to 2023 focused on selecting digital resources. *Journal of Computers in Education*. <https://doi.org/10.1007/s40692-025-00379-6>
- Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, 26(2), 91–108. <https://doi.org/10.1111/j.1471-1842.2009.00848.x>
- Grunert O'Brien, J., Millis, B. J., & Cohen, M. W. (2008). *The course syllabus: A learning-centered approach* (2nd ed.). Jossey-Bass.
- Harnish, R. J., & Bridges, K. R. (2011). Effect of syllabus tone: Students' perceptions of instructor and course. *Social Psychology of Education*, 14(3), 319–330. <https://doi.org/10.1007/s11218-011-9152-4>
- Judson, H. (2025). “But I didn't use ChatGPT!”: Democratic course design and generative artificial intelligence in higher education landscape. *Teaching Sociology*, 53(3), 246–256. <https://doi.org/10.1177/0092055X251342530>
- Khodabandelou, R., Yousef, A. M. F., & Elhaj, A. (2025). Affordances and core features of successful digital courses design in higher education. *Open Praxis*, 17(3), 544–560. <https://doi.org/10.55982/openpraxis.17.3.890>
- Kholid, M. N., Hendriyanto, A., Sahara, S., Muhaimin, L. H., Juandi, D., Sujadi, I., Kuncoro, K. S., & Adnan, M. (2023). A systematic literature review of Technological, Pedagogical and Content Knowledge (TPACK) in mathematics education: Future challenges for educational practice and research. *Cogent Education*, 10(2), 2269047. <https://doi.org/10.1080/2331186X.2023.2269047>
- Kurakbayeva, A., & Xembayeva, S. (2025). Enhancing professional abilities of university students through digital educational interventions: A study in Kazakhstani universities. *Frontiers in Education*, 9, 1478622. <https://doi.org/10.3389/feduc.2024.1478622>
- Mandulak, K. C. (2024). Implementing transparency in learning and teaching in higher ed in speech science coursework: A tutorial. *Perspectives of the ASHA Special Interest Groups*, 9(6), 1836–1852. https://doi.org/10.1044/2024_PERSP-24-00046
- Mazhit, Z., Amanzhol, M., Zhumagaliyeva, G., & Smagulova, M. (2023). Effectiveness of blended learning in higher education: Case of Astana IT University. *Bulletin of the Karaganda University Pedagogy Series*, 111(3), 180–185. <https://doi.org/10.31489/2023Ped3/180-185>
- Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017–1054. <https://doi.org/10.1111/j.1467-9620.2006.00684.x>
- Narbaev, T., Amirbekova, D., & Bakdaulet, A. (2025). A decade of transformation in higher education and science in Kazakhstan: A literature and scientometric review of national projects and research trends. *Publications*, 13(3), 35. <https://doi.org/10.3390/publications13030035>
- Nilson, L. B. (2007). *The graphic syllabus and the outcomes map: Communicating your course*. Jossey-Bass.
- Nurtayeva, D., Kredina, A., Kireyeva, A., Satybaldin, A., & Ainakul, N. (2024). The role of digital technologies in higher education institutions: The case of Kazakhstan. *Problems and Perspectives in Management*, 22(1), 562–577. [https://doi.org/10.21511/ppm.22\(1\).2024.45](https://doi.org/10.21511/ppm.22(1).2024.45)
- Palmer, M. S., Bach, D. J., & Streifer, A. C. (2014). Measuring the promise: A learning-focused syllabus rubric. *To Improve the Academy*, 33(1), 14–36. <https://doi.org/10.1002/tia2.20004>
- Parkes, J., & Harris, M. B. (2002). The purposes of a syllabus. *College Teaching*, 50(2), 55–61. <https://doi.org/10.1080/87567550209595875>
- Pereira, E., Nsair, S., Radin Pereira, L., & Grant, K. A. (2024). Constructive alignment in a graduate-level project management course: An innovative framework using large language models. *International Journal of Educational Technology in Higher Education*, 21(1), 30. <https://doi.org/10.1186/s41239-024-00457-2>
- Serbati, A., Picasso, F., Doria, B., & Grion, V. (2022). Learning outcomes and constructive alignment in the Mega-Universities syllabi: Which “promises” to students? *Form@re – Open Journal per la Formazione in Rete*, 22(2), 61–77. <https://doi.org/10.36253/form-13022>
- Slaterry, J. M., & Carlson, J. F. (2005). Preparing an effective syllabus: Current best practices. *College*

- Teaching, 53(4), 159–164. <https://doi.org/10.3200/CTCH.53.4.159-164>
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333–339. <https://doi.org/10.1016/j.jbusres.2019.07.039>
- Sulaiman, T. T. (2024). A systematic review on factors influencing learning management system usage in Arab gulf countries. *Education and Information Technologies*, 29(2), 2503–2521. <https://doi.org/10.1007/s10639-023-11936-w>
- Torraco, R. J. (2005). Writing integrative literature reviews: Guidelines and examples. *Human Resource Development Review*, 4(3), 356–367. <https://doi.org/10.1177/1534484305278283>
- Turnbull, D., Chugh, R., & Luck, J. (2022). An overview of the common elements of learning management system policies in higher education institutions. *TechTrends*, 66(5), 855–867. <https://doi.org/10.1007/s11528-022-00752-7>
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425–478. <https://doi.org/10.2307/30036540>
- Weimer, M. (2013). *Learner-centered teaching: Five key changes to practice* (2nd ed.). Jossey-Bass.
- Winkelmes, M.-A. (2013). Transparency in teaching: Faculty share data and improve students' learning. *Liberal Education*, 99(2), 48–55.
- Winkelmes, M.-A., Bernacki, M., Butler, J., Zochowski, M., Golanics, J., & Weavil, K. (2016). A teaching intervention that increases underserved college students' success. *Peer Review*, 18(1–2), 31–36.
- Zhanarstanova, M., Alimzhanov, E., Dossymkhan, Y., & Kaldybekova, A. (2024). The role of learning analytics in assessing the effectiveness of blended education: A case study of Astana IT University. In 2024 IEEE 4th International Conference on Smart Information Systems and Technologies (SIST) (pp. 447–454). IEEE. <https://doi.org/10.1109/SIST61555.2024.10629415>
- Zhubabayeva, T., Baumuratova, D., Zholshiyeva, L., Karabay, A., & Abdrakhmanov, K. (2025). Digital transformation in higher education: Toward a national model of digital university in Kazakhstan through global and local comparison. *Sustainability*, 17(24), 11132. <https://doi.org/10.3390/su172411132>
- Gauthier, G., Banner, J., & Winer, L. (2026). What is the syllabus for? Revealing tensions through a scoping review of syllabus uses. *International Journal for Academic Development*, 31(3), 479–498. <https://doi.org/10.1080/1360144X.2025.2465363>
- Fischer C, McPartlan P, Orona GA, Yu R, Xu D, Warschauer M. Salient syllabi: Examining design characteristics of science online courses in higher education. *PLoS One*. 2022 Nov 3;17(11):e0276839. doi: 10.1371/journal.pone.0276839. PMID: 36327299; PMCID: PMC9632807.

Information about authors:

Zhibek Tleshova – Candidate of Pedagogical Sciences, Associate Professor, Director of the School of General Education Disciplines, Astana IT University, Astana, Republic of Kazakhstan, zhibek.tleshova@astanait.edu.kz, ORCID: 0000-0001-5095-5436 (*corresponding author*).

Nurgissa Kabarbek – Astana IT University, Astana, Republic of Kazakhstan, 255390@astanait.edu.kz, ORCID: 0009-0002-9961-7176.

Bexultan Turgunov – Astana IT University, Astana, Republic of Kazakhstan, b.turgunov@astanait.edu.kz, ORCID: 0009-0002-8028-4359.

Received: 20.05.2025

Revised: 14.08.2026

Accepted: 28.09.2026

Published: 30.09.2026

Type: Review Article

¹Zhanar Bekbayeva, ²Moldir Seidina, ³Saule Aldabergenova, ⁴Botagoz Bermukhambetova*,
⁵Pawel Bogacz

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

² Kazakh National University of Sports (KNUS), Astana, Republic of Kazakhstan

³ Saken Seifullin Kazakh Agro Technical Research University (KazATU), Astana, Republic of Kazakhstan

⁴ National Center for Development Higher Education of the Ministry of Science and Higher Education of the Republic of Kazakhstan

⁴ AGH University of Krakow, Krakow, Republic of Poland

CRITICAL THINKING OF UNIVERSITY STUDENTS IN THE ERA OF ARTIFICIAL INTELLIGENCE

Abstract. The rapid expansion of artificial intelligence (AI) tools in higher education has reopened fundamental questions about the trajectory of students' critical thinking (CT) development. This article reports the results of a longitudinal cohort diagnostic study conducted at three higher education institutions in Astana, Republic of Kazakhstan L.N. Gumilyov Eurasian National University, Kazakh National University of Sports, and Saken Seifullin Kazakh Agro Technical Research University across four measurement points: 2020, 2022, 2024, and 2026, with a total sample of 334 students. CT was assessed using the Starkey Critical Thinking Test, adapted for Russian-language audiences, which measures analytical reasoning, argument evaluation, fallacy identification, evidence assessment, and problem-solving capacity. Inferential analyses employed the Mann–Whitney U test, Cohen's d effect sizes, and the Kruskal–Wallis omnibus test. Three research hypotheses were formulated and tested: that CT levels would demonstrate a statistically significant positive trajectory over the study period; that the period of active AI adoption (2023–2024) would coincide with a peak in CT performance; and that subsequent normalisation of AI use would be associated with a partial decline. The results confirm statistically significant improvement from 2020 ($M = 14.07$, $SD = 3.71$) to 2024 ($M = 18.58$, $SD = 4.44$), with a large effect size ($d = 1.10$, $p < 0.001$), and a statistically non-significant partial decline in 2026 ($M = 17.03$, $p = 0.059$). The overall Kruskal–Wallis test confirms significant between-cohort differences ($H = 43.13$, $p < 0.001$). The findings are discussed in relation to active pedagogical strategies, media literacy, and the dual role of AI as both a stimulus and a potential inhibitor of critical reasoning. Implications for competency-based curriculum reform within the Bologna Process framework are considered.

Keywords: critical thinking; higher education; longitudinal study; Starkey test; artificial intelligence; media literacy; active learning; Bologna Process; Kazakhstan; competency-based education.

Introduction

The question of whether university education genuinely develops students' capacity for critical, independent reasoning has acquired new urgency in the context of digitalisation and the widespread adoption of generative artificial intelligence tools. When AI systems can produce coherent, persuasive text on virtually any academic topic within seconds, the value added by higher education can no longer rest on information transmission alone. It must reside, above all, in the cultivation of the capacity to evaluate, question, and reason that is, in critical thinking (CT).

Kazakhstani higher education has been navigating this challenge within the broader context of Bologna Process-aligned reform. The State Programme for the Development of Education and Science of the Republic of Kazakhstan (2020–2025) explicitly designates critical thinking as one of the core broad-spectrum skills that university graduates must demonstrate, alongside research competence, ICT literacy, communicative flexibility, and problem-solving capacity (Government of the Republic of Kazakhstan, 2019). This programmatic commitment was deepened and extended by the Concept for the Development of Pre-School, Secondary, Technical and Vocational Education of the Republic of Kazakhstan for 2023–2029. The Concept reorients national educational policy from remedying systemic deficits toward placing the learner and their

future competencies at the centre, explicitly listing critical thinking, global competencies, and emotional intelligence as the defining outcomes of modernised curricula (Government of the Republic of Kazakhstan, 2023). In its revised edition, the Concept further stipulates the integration of artificial intelligence elements into state educational standards, with critical thinking and information literacy identified as the competencies to be cultivated through AI-enhanced learning environments. Yet the empirical evidence on whether CT is actually developing among students, and how that trajectory has been affected by the COVID-19 pandemic, post-pandemic pedagogical restructuring, and the emergence of generative AI, remains scarce particularly from Central Asian institutional contexts.

The present study directly addresses this gap. Drawing on a series of diagnostic measurements conducted at three higher education institutions in Astana: L.N. Gumilyov Eurasian National University (ENU), Kazakh National University of Sports (KNUS), and Saken Seifullin Kazakh Agro Technical Research University (KazATU), between 2020 and 2026, we report the first longitudinal cohort analysis of CT development spanning the pre-pandemic, post-pandemic, early AI, and normalised AI periods in this regional context. The study builds on the authors' earlier work on CT diagnostics and active learning strategies in Kazakhstani vocational and higher education (Bekbayeva et al., 2020; Bekbayeva et al., 2022; Bekbayeva & Seidina, 2025), extending it with new data and a comparative longitudinal framework.

Three research hypotheses structure the inquiry. H1: CT levels among students will demonstrate a statistically significant positive trajectory over the 2020–2026 period, reflecting cumulative effects of competency-oriented pedagogical reform. H2: The period coinciding with the initial adoption of generative AI tools (2023–2024) will mark a peak in CT performance, reflecting the stimulating effect of novel technological engagement on analytical reasoning. H3: Subsequent normalisation of AI use (2025–2026) will be associated with a partial, though not necessarily significant, decline in CT scores, reflecting a shift from active to passive AI engagement patterns.

Literature Review

Critical Thinking: Conceptual Foundations

Critical thinking in higher education has been conceptualised along three main dimensions in the contemporary literature. The cognitive dimension encompasses discrete skills including analysis, inference, evaluation, and argumentation that allow students to assess the quality of information and construct logically warranted conclusions (Yessekeshova et al., 2020; Bezanilla et al., 2021). The metacognitive dimension involves the conscious regulation of one's own reasoning processes: monitoring for bias, recognising gaps in understanding, and adjusting cognitive strategies in response to feedback (Rivas et al., 2022). The dispositional dimension refers to stable intellectual orientations as curiosity, open-mindedness, and epistemic humility that determine whether a student who possesses CT skills actually applies them spontaneously in novel contexts (Álvarez-Huerta et al., 2023).

This three-part framework is particularly significant for the Kazakhstani context, where educational reform has traditionally emphasised knowledge acquisition over the cultivation of evaluative and reflective capacities. The competency-based approach introduced through Bologna Process alignment requires that graduates not merely know content but demonstrate the ability to apply knowledge critically, handle uncertainty, and make informed judgements, all of which presuppose well-developed CT in all three dimensions (Temirgazina et al., 2024).

Factors Shaping Critical Thinking Development

The empirical literature identifies pedagogical, psychological, sociocultural, and technological factors as the principal determinants of CT outcomes in university students. Among pedagogical factors, the most robustly evidenced approaches include problem-based learning (PBL), multi-level task design, inquiry-based learning, and case-based discussion (Yu & Zin, 2023; Bekbayeva et al., 2022). Multi-level task structures in particular where the cognitive demand of tasks is calibrated to challenge students progressively across analysis, synthesis, and evaluation have been shown to produce measurable CT gains among university students (Bekbayeva et al., 2021).

Media resources constitute an additional pedagogical lever whose potential for CT development has received growing attention. The use of audio-visual materials, digital texts, and media content as objects of critical analysis requiring students to evaluate source credibility, identify persuasive strategies, and assess

evidence quality engages multiple CT skill dimensions simultaneously. Bekbayeva and Seidina (2025), in a study of 218 students at a Kazakhstani university using a multi-level media task design, demonstrated statistically significant CT improvements (Cramer–Welch test), confirming the effectiveness of this approach for developing analytical, synthesis, and evaluation competencies in Russian-language academic contexts.

Among psychological factors, metacognition has emerged as the most consistent predictor of CT growth. Rivas, Saiz, and Ossa (2022) demonstrated through controlled experimentation that metacognitive self-regulation training produces substantially larger CT improvements than skills-focused instruction alone, establishing metacognition as a causal mechanism rather than merely a correlate of CT performance. Guamanga, Saiz, Rivas, and Morales Bueno (2025) further showed, using structural equation modelling, that CT predicts metacognition, which in turn predicts psychological well-being and empathy positioning CT as a core personal quality with implications beyond academic performance.

Technological factors present a more complex and contested picture. The systematic review by Melisa et al. (2025), covering AI's effects on CT in higher education, identified a paradoxical dynamic: students who engage critically with AI-generated content as questioning, fact-checking, and evaluating it tend to develop stronger CT skills, while those who use AI primarily to replace independent analytical effort show diminished reasoning capacity over time. This paradox is further elaborated in the bibliometric analysis of Salido et al. (2025), which mapped five thematic clusters in the AI-and-CT literature and identified pedagogical innovation and AI literacy as the two areas of most rapid growth since the widespread adoption of ChatGPT in 2023.

CT Diagnostics: Instruments and Approaches

The assessment of CT in empirical research relies on a combination of standardised psychometric instruments and performance-based tasks. Among internationally validated measures, the Watson-Glaser Critical Thinking Appraisal, the California Critical Thinking Skills Test, and the PENCRIASAL (developed by Rivas and Saiz at the University of Salamanca) are most widely used in Scopus- and WoS-indexed research (Araújo et al., 2024; Wang & Abdullah, 2024). A critical limitation of these instruments in the Central Asian context is the absence of validated Russian-language adaptations with established norms for Kazakhstani university populations.

The Starkey Critical Thinking Test, adapted for Russian-language audiences and employed in the present study, addresses this gap. It is a 27-item multiple-choice instrument assessing applied reasoning across five competency areas: argument recognition and evaluation, fallacy identification, evidence assessment, source credibility evaluation, and solution analysis. Its psychometric properties in the Kazakhstani context have been established in earlier work by the research team (Bekbayeva & Seidina, 2025; Bekbayeva et al., 2020), providing the methodological continuity necessary for longitudinal comparison. Score interpretation follows three levels: Low (0–9 points), Medium (10–18 points), and High (19–27 points).

Research Gap and Contribution

Despite significant international activity, CT research from Central Asian higher education contexts remains substantially underrepresented in high-impact indexed publications (Andreucci-Annunziata et al., 2023). The present study addresses this gap in three ways: by providing longitudinal diagnostic data spanning six years and four distinct technological-pedagogical phases; by employing a locally validated instrument with established comparability across measurement points; and by situating empirical findings within the theoretical and policy frameworks most relevant to Kazakhstani higher education as the Bologna Process competency agenda and the national education development programme.

Methodology

Research Design and Rationale

The study employs a longitudinal cohort design with four independent cross-sectional measurement points: October 2020, May 2022, December 2024, and March 2026. This design comparing distinct student cohorts at successive timepoints rather than tracking the same individuals allows examination of how CT levels among incoming university students have shifted across successive academic generations, each encountering a different technological and pedagogical environment. While individual-level longitudinal tracking would permit stronger causal claims, the cohort design is appropriate for the research objectives and consistent with practice in educational trend research (Wang & Abdullah, 2024).

The four measurement points correspond to four analytically distinct contextual phases: the pre-pandemic

baseline (2020); the early post-pandemic period, marked by rapid transition to blended learning (2022); the first year of widespread generative AI availability in Kazakhstan (2024); and the period of normalised AI integration (2026). This temporal structure makes the dataset uniquely suited to examining the hypothesised relationships between CT development and technological context.

Participants

The total sample comprised 334 students across four cohorts: 2020 ($n = 91$; female = 54, male = 37), 2022 ($n = 48$), 2024 ($n = 59$; female = 32, male = 27), and 2026 ($n = 136$; female = 116, male = 20). Participants were recruited from three higher education institutions in Astana, Republic of Kazakhstan: L.N. Gumilyov Eurasian National University, the Kazakh National University of Sports (KNUS), and Saken Seifullin Kazakh Agro Technical Research University (KazATU). All three institutions have been implementing Bologna Process-aligned competency-based curricula, making the multi-institutional sample representative of the broader Kazakhstani higher education reform context. Students were enrolled in first and second years of undergraduate programmes spanning agricultural, natural science, and pedagogical fields. The age range was predominantly 18–21 years. Participation was entirely voluntary; students were informed of the study objectives and guaranteed full confidentiality of their individual results. The notable gender imbalance in the 2026 cohort (85.3% female) reflects programme-specific enrolment patterns at the participating institutions in that academic year and is acknowledged as a potential confounding variable in the interpretation of results.

Instrument

The Starkey Critical Thinking Test is a 27-item multiple-choice instrument assessing applied reasoning in realistic everyday and academic scenarios. Each item presents a situation and requires the student to identify the most logically defensible response from four alternatives. Items target five competency areas: argument recognition and evaluation (items 1–6); fallacy identification, including ad hominem, post hoc reasoning, false dichotomy, and circular reasoning (items 7–13); evidence quality assessment (items 14–18); source credibility evaluation (items 19–22); and problem identification and solution analysis (items 23–27). Correct responses are scored 1 point each; total scores range from 0 to 27. The test was administered digitally through a Google Forms interface, with individual responses and total scores recorded automatically. The Russian-language adaptation used in this study has been applied consistently across all four measurement points, ensuring instrument comparability (Bekbayeva & Seidina, 2025).

Statistical Analysis

Data were processed using Python 3.11 (scipy.stats library). For each cohort, descriptive statistics were computed including mean (M), standard deviation (SD), standard error (SE), median, minimum, and maximum. Distribution normality was assessed using the Shapiro–Wilk test. Since the 2024 ($W = 0.944$, $p = 0.009$) and 2026 ($W = 0.973$, $p = 0.008$) cohorts violated normality assumptions, all inferential comparisons used non-parametric methods. Pairwise between-cohort comparisons employed the Mann–Whitney U test (two-tailed, $\alpha = 0.05$). Effect sizes were computed as Cohen's d using pooled standard deviations, with conventional thresholds of small (0.2–0.5), medium (0.5–0.8), and large (> 0.8) following Cohen (1988). Between-group differences across all four cohorts were assessed using the Kruskal–Wallis H test.

Results And Discussion

Descriptive Statistics and Level Distribution

Table 1 presents the combined descriptive statistics and CT level distributions for all four cohorts. The data reveal a pronounced non-linear trajectory: an upward trend from 2020 to 2024, a peak in 2024, and a partial reversal in 2026. The mean score rose from 14.07 in 2020 to 18.58 in 2024 — an increase of 4.51 points, corresponding to a shift from the middle of the Medium level to the threshold of the High level. In 2026, the mean declined to 17.03, remaining above the 2022 value but below the 2024 peak.

Table 1*Descriptive statistics and CT level distribution by cohort (Starkey Critical Thinking Test, 0–27 points)*

Year	N	M	SD	Min	Max	Mdn	SE	Low, %	Med, %	High, %
2020	91	14.07	3.71	6	22	14.0	0.39	12.1	74.7	13.2
2022	48	15.42	3.61	7	22	16.0	0.52	4.2	75.0	20.8
2024	59	18.58	4.44	7	25	19.0	0.58	1.7	39.0	59.3
2026	136	17.03	5.28	2	27	17.0	0.45	10.3	47.8	41.9

The level distribution data amplify this picture. The proportion of students achieving the High CT level rose dramatically from 13.2% in 2020 to 59.3% in 2024 — a more than fourfold increase over four years. Simultaneously, the Low level proportion fell from 12.1% to 1.7%. In 2026, the High level proportion declined to 41.9% while the Low level partially recovered to 10.3%, suggesting a redistribution of students from the High level toward the Medium and Low categories. The standard deviation also increased from 3.71 (2020) to 5.28 (2026), indicating growing heterogeneity in CT outcomes — a pattern consistent with diverging student engagement patterns under conditions of widespread AI availability.

Hypothesis Testing: Inferential Results

The omnibus Kruskal–Wallis test confirmed statistically significant overall differences across all four cohorts ($H = 43.13$, $p < 0.001$), warranting pairwise comparisons. Table 2 presents all six pairwise Mann–Whitney U comparisons with associated statistics.

Table 2*Pairwise Mann–Whitney U test results with effect sizes ($\alpha = 0.05$)*

Comparison	U	p	Sig.	Cohen's d	Effect	ΔM	Direction
2020 vs 2022	1730	0.044	*	0.37	Small	+1.35	↑
2020 vs 2024	1176	<0.001	***	1.10	Large	+4.51	↑
2020 vs 2026	3837	<0.001	***	0.65	Medium	+2.96	↑
2022 vs 2024	802	<0.001	***	0.78	Medium	+3.16	↑
2022 vs 2026	2485	0.014	*	0.36	Small	+1.61	↑
2024 vs 2026	4694	0.059	n.s.	0.32	Small	–1.55	↓

Note. Sig. = significance level: *** $p < 0.001$; * $p < 0.05$; n.s. = not significant. ΔM = difference in group means (positive = later group higher).

Hypothesis H1: that CT levels would demonstrate a statistically significant positive trajectory, is supported. All comparisons involving the 2020 baseline cohort reached statistical significance. The comparison between 2020 and 2024 produced the largest observed effect in the dataset ($d = 1.10$), qualifying as large by conventional benchmarks and representing the most robust evidence of CT improvement across the study period. Importantly, even the comparison between 2020 and 2026 spanning the full six-year range remains highly significant ($p < 0.001$, $d = 0.65$, medium effect), confirming that CT gains were not entirely reversed by 2026.

Hypothesis H2: that the 2024 cohort would represent a peak in CT performance, is also supported. The 2024 cohort recorded the highest mean score (18.58) and the largest proportion of High-level students (59.3%) across all measurement points. The comparison between 2022 and 2024 yielded a medium effect ($d = 0.78$, $p < 0.001$), confirming that 2024 performance was substantially superior to the previous measurement point, not merely to the 2020 baseline.

Hypothesis H3: that the 2026 cohort would show a partial decline, receives qualified support. The comparison between 2024 and 2026 shows a mean difference of –1.55 points in the expected direction, but this difference does not reach statistical significance ($p = 0.059$). Strictly interpreted, H3 cannot be confirmed at the 0.05 level. However, the p-value of 0.059 represents a borderline result, and the observed decline is substantively meaningful, particularly when combined with the structural changes in level distribution (High

level falling from 59.3% to 41.9%; Low level recovering from 1.7% to 10.3%). This pattern warrants continued monitoring and does not permit dismissal of the underlying dynamic.

Interpreting the Longitudinal Trajectory

The overall pattern presents significant rise to 2024, borderline non-significant partial reversal in 2026 and admits several interpretations that are not mutually exclusive and may operate simultaneously.

The first interpretation centres on pedagogical restructuring. The period 2020–2024 was characterised by rapid institutional adaptation to online and blended learning, driven by pandemic necessity. This transition compelled faculty to redesign courses around interactive, problem-oriented formats, precisely the active learning approaches that the literature consistently identifies as most effective for CT development (Yu & Zin, 2023; Bekbayeva et al., 2022). The introduction of media-based multi-level tasks and structured critical analysis activities approaches empirically validated in the Kazakhstani context (Bekbayeva & Seidina, 2025) likely contributed directly to the observed improvements. The partial decline in 2026 may reflect a reversion toward conventional instructional formats as the emergency stimulus of the pandemic receded.

The second interpretation, which we term the AI novelty hypothesis, draws on the paradox identified in the systematic review literature (Melisa et al., 2025; Salido et al., 2025). When ChatGPT and similar tools became widely accessible to Kazakhstani students in early 2023, the novelty of the technology prompted active engagement: students interrogated AI outputs, compared them with other sources, questioned their accuracy, and discussed their limitations. This metacognitive engagement with novel AI content may have inadvertently stimulated the same analytical processes that CT instruction aims to cultivate. By 2026, however, two years of normalised AI use may have shifted many students toward more passive consumption patterns — accepting AI outputs without critical evaluation, delegating analytical tasks to AI systems, and thereby reducing the frequency of independent reasoning exercises.

The third interpretation addresses sampling variability. The 2026 cohort was substantially larger ($n = 136$) and exhibited a markedly different gender composition (85.3% female) compared with earlier cohorts. If gender is associated with differences in CT performance in the local context, a question that was not the focus of this study and cannot be resolved from available data, this compositional shift could partially account for the lower mean score. The wider standard deviation in 2026 (5.28 vs. 4.44 in 2024) is also consistent with a more heterogeneous intake, potentially reflecting broader programme-level enrolment patterns rather than a systematic CT decline.

The fourth interpretation posits stabilisation rather than decline: that the sustainable CT equilibrium under current institutional conditions lies in the range of 16–18 points, with the 2024 peak reflecting a transient confluence of favourable conditions: intensive pedagogical restructuring, technology novelty effect, and heightened student motivation, unlikely to be replicated without deliberate and sustained intervention. This interpretation is consistent with the non-significant p-value for the 2024–2026 comparison and suggests that the primary research task is not to explain a decline but to identify the conditions that enabled the 2024 peak and determine how to sustain them.

Across all four interpretations, a common implication emerges: CT development in higher education is neither automatic nor irreversible. It depends on active pedagogical investment, intentional engagement with analytical tasks, and the cultivation of metacognitive habits that enable students to monitor and improve their own reasoning. Psychological and pedagogical guidance approaches that explicitly address these dimensions including coaching, reflective practice, and deliberate AI-critical literacy training offer a promising pathway (Phembayev & Uaikhanova, 2024). The competency model of a Kazakhstani pedagogical university graduate, as defined in state regulatory documents, provides the normative framework within which such approaches should be embedded (Temirgazina et al., 2024).

Implications for Policy and Practice

The findings carry several concrete implications for Kazakhstani higher education institutions operating within the Bologna Process framework. First, the significant positive trajectory from 2020 to 2024 demonstrates that CT improvement is achievable at the institutional level within meaningful timeframes, provided that pedagogical conditions actively support analytical reasoning. This should encourage investment in active learning infrastructure, faculty development in CT pedagogy, and the systematic integration of multi-level and media-analytical task designs across disciplinary curricula.

Second, the borderline decline in 2026, combined with the structural changes in level distribution, signals

the need for proactive AI literacy integration into CT pedagogy. Rather than treating AI as a threat to student reasoning capacity, institutions should design learning environments in which AI outputs serve as objects of critical analysis as stimuli for evaluation, comparison, and argumentation, rather than as substitutes for independent thought. This pedagogical reframing of AI aligns directly with the CT competency goals of the State Programme for the Development of Education and Science.

Third, the growing heterogeneity of CT outcomes (reflected in widening standard deviations) suggests that average scores increasingly mask divergent student trajectories. Diagnostic monitoring should be accompanied by differentiated pedagogical support: targeted intervention for students at the Low level, challenge-based extension for those at the High level, and structured metacognitive scaffolding for the majority in the Medium range.

Conclusion

This longitudinal cohort study of critical thinking among university students in Kazakhstan, spanning four measurement points from 2020 to 2026 and involving 334 participants, yields findings of both theoretical and practical significance. The data confirm a substantial and statistically robust improvement in CT from the pre-pandemic baseline to the period of initial AI adoption, with the 2024 cohort recording the highest mean score (18.58 out of 27) and the largest proportion of High-level students (59.3%) across the entire observation period. A borderline-significant partial decline in 2026 raises important questions about the sustainability of CT gains under conditions of normalised AI use, though the 2026 cohort still significantly outperforms the 2020 baseline.

Three hypotheses were formulated and tested. H1, predicting a significant positive trajectory over the study period, is fully supported. H2, predicting that the 2024 cohort would represent a performance peak, is supported. H3, predicting a partial decline in 2026, receives qualified support: the decline is substantively meaningful and in the predicted direction, but does not reach conventional statistical significance at the 0.05 threshold.

The study contributes original empirical evidence from a Kazakhstani higher education context that has been underrepresented in international CT research. By employing a locally validated instrument (the Starkey Critical Thinking Test in its Russian-language adaptation) consistently across all measurement points, and by situating findings within both the theoretical literature on CT and the policy framework of Kazakhstani educational reform, the study offers a model for systematic longitudinal CT monitoring at the institutional level.

The limitations of the study are acknowledged: the cross-sectional cohort design precludes individual-level causal inference; the geographically concentrated multi-institutional sample limits transferability; the absence of AI usage data prevents direct testing of the AI paradox hypothesis; and the gender imbalance in the 2026 cohort introduces potential confounding. Future research should address these limitations through longitudinal individual tracking, multi-institutional sampling, integration of AI engagement pattern measurement, and gender-stratified analysis. The ongoing research programme of the present authors, building on this diagnostic foundation, will extend the analysis to include intervention studies and qualitative exploration of student reasoning processes.

Conflict Of Interest Statement

The authors declare no potential conflicts of interest with respect to the research, authorship, or publication of this article.

Author Contributions

Pawel Bogacz: International Context and Comparative Perspective, Contribution of Information on International Students, Review and Editing. Zhanar Bekbayeva: Writing – Original Draft Preparation, Development of the Main Section of the Article. Moldir Seidina: Investigation, Data Collection and Research Conducted at Her University, Review and Editing. Saule Aldabergenova: Investigation, Data Collection and Research Conducted at Her University, Review and Editing. Botagoz Bermukhambetova: Article Formatting, Investigation in the Area of Digital Competence Development, Writing – Review and Editing.

Declaration of Use of Generative AI and AI-Assisted Technologies

The authors declare that no generative AI or AI-assisted technologies were used in the preparation of this manuscript.

References

- Álvarez-Huerta, P., Muela, A., & Larrea, I. (2023). Disposition towards critical thinking and student engagement in higher education. *Innovative Higher Education*, 48(2), 239–256. <https://doi.org/10.1007/s10755-022-09614-9>
- Andreucci-Annunziata, P., Riedemann, A., Cortés, S., Mellado, A., del Río, M. T., & Vega-Muñoz, A. (2023). Conceptualizations and instructional strategies on critical thinking in higher education: A systematic review of systematic reviews. *Frontiers in Education*, 8, 1141686. <https://doi.org/10.3389/educ.2023.1141686>
- Araújo, B., Gomes, S. F., & Ribeiro, L. (2024). Critical thinking pedagogical practices in medical education: A systematic review. *Frontiers in Medicine*, 11, 1358444. <https://doi.org/10.3389/fmed.2024.1358444>
- Bekbayeva, Z. S., & Seidina, M. Z. (2025). Using media resources in education to develop university students' critical thinking skills. *ILIM*, 43(1), 8–17. <https://doi.org/10.47751/skpu.1937.v43i1.1>
- Bekbayeva, Z. S., Solovyev, A. N., Chechetkina, N. V., Bolshakov, A. A., & Sosina, L. V. (2022). Problems of organizing critical thinking training for university students. *Journal of Higher Education Theory and Practice*, 22(10), 116–123. <https://doi.org/10.33423/jhetp.v22i10.5390>
- Bekbayeva, Z. S., Galiyev, T. T., Albytova, N., Zhazykbayeva, Z. M., & Mussatayeva, A. B. (2021). Fostering post-secondary vocational students' critical thinking through multi-level tasks in learning environments. *World Journal on Educational Technology: Current Issues*, 13(3), 397–406. <https://doi.org/10.18844/wjet.v13i3.5948>
- Bezanilla, M. J., Fernández-Nogueira, D., Poblete, M., & Galindo-Domínguez, H. (2021). Methodologies for teaching-learning critical thinking in higher education: The teacher's view. *Thinking Skills and Creativity*, 42, 100892. <https://doi.org/10.1016/j.tsc.2021.100892>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Government of the Republic of Kazakhstan. (2019). State Programme for the Development of Education and Science of the Republic of Kazakhstan for 2020–2025 [Государственная программа развития образования и науки Республики Казахстан на 2020–2025 годы]: Government Resolution No. 988. Expired upon completion of implementation cycle. <https://www.gov.kz/memleket/entities/edu/documents/details/14703?lang=ru>
- Government of the Republic of Kazakhstan. (2023). Concept for the Development of Pre-School, Secondary, Technical and Vocational Education of the Republic of Kazakhstan for 2023–2029 [Концепция развития дошкольного, среднего, технического и профессионального образования Республики Казахстан на 2023–2029 годы]: Government Resolution No. 249 of 28 March 2023 (as amended by Resolution No. 465 of 13 June 2024; currently in force). <https://adilet.zan.kz/rus/docs/P2300000249>
- Guamanga, M. H., Saiz, C., Rivas, S. F., & Morales Bueno, P. (2025). Critical thinking and metacognition: Pathways to empathy and psychological well-being. *Journal of Intelligence*, 13(3), 34. <https://doi.org/10.3390/jintelligence13030034>
- Jaramillo Gómez, D. L., Álvarez Maestre, A. J., Parada Trujillo, A. E., Pérez Fuentes, C. A., Bedoya Ortiz, D. H., & Sanabria Alarcón, R. K. (2025). Determining factors for the development of critical thinking in higher education. *Journal of Intelligence*, 13(6), 59. <https://doi.org/10.3390/jintelligence13060059>
- Melisa, R., Ashadi, A., Triastuti, A., Hidayati, S., Salido, A., Ero, P. E. L., Marlini, C., Zefrin, Z., & Al Fuad, Z. (2025). Critical thinking in the age of AI: A systematic review of AI's effects on higher education. *Educational Process: International Journal*, 14, e2025031. <https://doi.org/10.22521/edupij.2025.14.31>
- Phembayev, M. A., & Uaikhanova, M. (2024). Comparative analysis of foreign experience in implementing psychological and pedagogical guidance for soft skills development of university students. *Higher Education in Kazakhstan*, 47(3), 79–86. <https://doi.org/10.59787/2413-5488-2024-47-3-79-86>

- Rivas, S. F., Saiz, C., & Ossa, C. (2022). Metacognitive strategies and development of critical thinking in higher education. *Frontiers in Psychology*, 13, 913219. <https://doi.org/10.3389/fpsyg.2022.913219>
- Salido, A., Ero, P. E. L., Melisa, R., Marlini, C., Ashadi, A., Triastuti, A., Hidayati, S., Zefrin, Z., & Al Fuad, Z. (2025). Integrating critical thinking and artificial intelligence in higher education: A bibliometric and systematic review of skills and strategies. *Social Sciences & Humanities Open*, 12, 101924. <https://doi.org/10.1016/j.ssaho.2025.101924>
- Temirgazina, Z., Andryuchshenko, O., & Sirlibaev, M. (2024). Reflection of the competence model of a pedagogical university graduate in regulative documents. *Higher Education in Kazakhstan*, 46(2), 118–127. <https://doi.org/10.59787/2413-5488-2024-46-2-118-127>
- Wang, Q., & Abdullah, A. H. (2024). Enhancing students' critical thinking through mathematics in higher education: A systematic review. *SAGE Open*, 14(3). <https://doi.org/10.1177/21582440241275651>
- Yessekeshova, M. D., Ibrayeva, K. Zh., Kaltayeva, G. K., Albytova, N., & Bekbayeva, Z. S. (2020). On the formation of critical thinking of students of a higher educational institution. *Journal of Intellectual Disability – Diagnosis and Treatment*, 8(3), 388–395. <https://doi.org/10.6000/2292-2598.2020.08.03.15>
- Yu, L., & Zin, Z. M. (2023). The critical thinking-oriented adaptations of problem-based learning models: A systematic review. *Frontiers in Education*, 8, 1139987. <https://doi.org/10.3389/feduc.2023.1139987>

Information about authors

Zhanar Bekbayeva – PhD, Senior Lecturer, Department of Pedagogy, L.N. Gumilyov Eurasian National University, Astana, Republic of Kazakhstan; e-mail: bekbayeva@inbox.ru; ORCID: 0000-0002-6070-4195.

Moldir Seidina – PhD, Associate Professor, Department of Management and Innovation in Sports, Kazakh National University of Sports, Astana, Republic of Kazakhstan; e-mail: m_seidina@apems.edu.kz; ORCID: 0000-0002-9106-9147.

Saule Aldabergenova – PhD, Head of Educational Programs in Transport Engineering, Technology, Standardization, Certification, and Metrology, S. Seifullin Kazakh Agrotechnical Research University, Astana, Republic of Kazakhstan; e-mail: s.algabergenova@kazatu.edu.kz; ORCID: 0000-0001-6297-7632.

Botagoz Bermukhambetova – Deputy Director, Higher Education Development National Center of the Ministry of Science and Higher Education of the Republic of Kazakhstan, Astana, Republic of Kazakhstan; e-mail: b.bermukhambetova@n-k.kz; ORCID: 0000-0003-1414-6484. (**corresponding author**)

Pawel Bogacz – DSc, PhD Eng., Professor at AGH University of Krakow, Department of Economics and Management in Industry, Faculty of Civil Engineering and Resource Management, Krakow, Republic of Poland; e-mail: bogacz@agh.edu.pl; ORCID: 0000-0002-0099-6376.

Received: 02.07.2026

Revised: 31.08.2026

Accepted: 21.09.2026

Published: 30.09.2026

Type: Research Article

Damira Jantassova*, Eleonora Prozovskaya

Abylkas Saginov Karaganda Technical University, Republic of Kazakhstan, Karaganda

EXPLORING THE IMPACT OF AI CHATBOTS ON ENGLISH-SPEAKING SKILLS IN ENGINEERING EDUCATION: A MIXED-METHODS STUDY

Abstract. This study examines the effectiveness of AI-powered chatbots in developing English-speaking skills among Pre-Intermediate engineering students. The relevance of the research is driven by the need to expand opportunities for oral language practice in technical universities, where traditional instructional approaches often fail to provide sufficient communicative engagement for learners. The study employed a mixed-methods research design and was conducted at Abylkas Saginov Karaganda Technical University. The pedagogical intervention focused on enhancing students' speaking proficiency through the use of interactive AI chatbots, incorporating role-play conversations, spontaneous speaking practice, and voice-based exercises. The findings revealed a substantial improvement in the performance of the experimental group, including increased speaking fluency, greater grammatical accuracy, and reduced levels of communication anxiety. The analysis of the collected data further demonstrated enhanced linguistic self-confidence, a lower fear of making mistakes, and increased learner autonomy. At the same time, several limitations of AI-based tools were identified, including occasional inaccuracies in interpreting user input and a restricted capacity to sustain in-depth conversations. Overall, the results confirm the pedagogical value of AI chatbots as an effective instrument for fostering oral English-language communication skills within the context of technical higher education.

Keywords: artificial intelligence, AI chatbots, speaking skills, engineering education, mixed-methods research, language anxiety, digital pedagogy.

Introduction

The rapid advancement of artificial intelligence (AI) technologies is exerting a profound influence on higher education, creating new opportunities for personalized learning, enhanced interactivity, and improved instructional effectiveness (Winkler & Söllner, 2018). Among the most promising AI-driven tools are chatbots capable of sustaining natural, real-time communication with users. Owing to their ability to simulate authentic dialogue, provide immediate feedback, and adapt interactions to learners' individual needs, AI chatbots are increasingly being used to support the development of English-speaking skills in educational settings (Gutiérrez, 2023).

The integration of AI chatbots is particularly relevant in the context of engineering education. In today's professional environment, engineers are expected not only to possess solid technical expertise but also to demonstrate effective communication skills in English. Such competencies are essential for participation in international projects, professional discussions, academic mobility programmes, and collaboration with colleagues from different countries. Nevertheless, the development of English-speaking skills among engineering students remains one of the most challenging objectives of language education. The nature of study at technical universities requires students to devote a substantial portion of their academic time to discipline-specific courses, which often limits opportunities for systematic oral language practice. In addition, many engineering students experience difficulties in engaging in spontaneous communication in English, demonstrate low confidence during oral interactions, and experience high levels of communication anxiety. As a result, even learners who possess an adequate vocabulary and a sound knowledge of grammatical structures frequently struggle to achieve speaking fluency and sustain meaningful communication in English.

Within this context, AI chatbots represent a promising tool for the development of speaking skills. Unlike traditional approaches to organizing oral language practice, chatbots provide virtually unlimited opportunities for communicative interaction, enable learners to receive immediate feedback, and facilitate language practice beyond the classroom setting. Another important advantage is the creation of a psychologically supportive environment in which the fear of making mistakes is reduced and students become

more willing to use English for communication. Furthermore, the integration of such tools into learners' familiar digital environment may enhance both their engagement and motivation in the process of learning English.

Despite the growing scholarly interest in the application of artificial intelligence in language education, most existing studies have focused primarily on the quantitative assessment of learning outcomes. However, a comprehensive understanding of the pedagogical potential of AI chatbots requires consideration not only of objective indicators of speaking-skill development but also of learners' subjective experiences, perceptions of the technology, attitudes toward its use, and the advantages and limitations identified during the learning process (Winkler & Söllner, 2018; Gutiérrez, 2023).

To address this gap, the present study adopts a mixed-methods research design that combines quantitative and qualitative approaches to data collection and analysis. The quantitative phase of the study seeks to evaluate changes in students' speaking proficiency following the integration of AI chatbots into English language instruction. The qualitative phase explores learners' experiences, their attitudes toward chatbot-assisted language learning, and their perceptions of the role of AI chatbots in fostering oral communication skills.

The study was conducted at Abylkas Saginov Karaganda Technical University and involved Pre-Intermediate students enrolled in engineering degree programmes. The primary aim of the study was to investigate the impact of AI chatbots on the development of English-speaking skills among engineering students and to examine learners' perceptions of chatbots as a language-learning tool through the application of a mixed-methods research approach.

Based on the identified research gap, the study hypothesized that the integration of AI chatbots into speaking instruction for students of technical specialties would result in a significant improvement in oral proficiency, a reduction in communicative anxiety, and an increase in learners' linguistic activity at the Pre-Intermediate level, compared with traditional instructional methods. This hypothesis was theoretically grounded in the pedagogical affordances of AI chatbots, particularly their capacity to facilitate interactive communication, provide immediate feedback, adapt to individual learner needs, and simulate authentic communicative situations. These affordances were expected to support a lower-anxiety learning environment, facilitate sustained opportunities for oral language practice, and foster greater learner autonomy in speaking activities.

Methodology and Research Materials

To provide a comprehensive evaluation of the effectiveness of AI chatbots in developing English-speaking skills, the study adopted a mixed-methods research approach. A Convergent Study Design (CSD) was selected as the primary research strategy, enabling the integration of quantitative and qualitative data and facilitating a deeper interpretation of the findings.

The research was conducted at Abylkas Saginov Karaganda Technical University. The sample consisted of engineering students studying English at the Pre-Intermediate level. All participants demonstrated a comparable level of language proficiency, as confirmed by the results of a placement test administered prior to the study.

The quantitative assessment focused on four key dimensions of oral proficiency: fluency, grammatical accuracy, lexical resource, and pronunciation. These dimensions were selected to provide a comprehensive assessment of students' speaking performance and to identify changes associated with the use of AI chatbots in English language instruction. The implementation of the convergent design involved two complementary methods of data collection.

The quantitative component consisted of pre-test and post-test assessments designed to measure changes in students' speaking performance. The assessments were administered before and after the intervention to ensure comparability of the results.

The qualitative component was conducted through semi-structured individual interviews with participants from the experimental group. This instrument enabled the exploration of students' experiences with AI chatbots, their perceptions of chatbot-assisted language learning, and the influence of AI-mediated interaction on communication anxiety, language confidence, and motivation for independent speaking practice. The interviews addressed several thematic areas, including overall experiences with AI chatbots, perceived changes in speaking fluency, pronunciation and vocabulary development, the impact of chatbot

interaction on confidence in oral communication, and the perceived advantages and limitations of AI tools in English language learning. The qualitative findings complemented the quantitative results and provided deeper insights into the mechanisms through which AI technologies may influence speaking-skill development among engineering students.

The triangulation of quantitative and qualitative methods enhanced the validity of the study by not only identifying measurable changes across the four speaking indicators but also revealing the underlying pedagogical mechanisms through which AI technologies affect language learning outcomes.

The research was conducted using a set of complementary pedagogical and scientific methods, including:

Theoretical analysis and systematization. A comprehensive analysis of contemporary language-teaching approaches to speaking instruction was conducted, with particular emphasis on communicative and interactive methodologies. Special attention was given to the psycholinguistic factors underlying the development of oral communication skills.

Literature review and analysis. Previous studies investigating the application of artificial intelligence and generative language models in education were examined. Particular attention was paid to research exploring the use of AI chatbots and automated feedback systems for developing speaking skills in English language-learning contexts (Gutiérrez, 2023; Beisembayeva, 2025; Khamsa & Zagidullina, 2025; Lobeeva, 2023).

Analysis of practical applications. Existing cases of integrating digital pedagogical tools into educational practice were examined. The functional capabilities of platforms such as ChatGPT Live, Google Gemini, Duolingo Max, Talkpal AI, Character.AI, Claude AI, and Microsoft Copilot were analyzed with regard to their potential for supporting speaking-skill development (Gutiérrez, 2023; Beisembayeva, 2025; Khamsa & Zagidullina, 2025).

Empirical investigation. The empirical phase of the study involved 80 engineering students studying English at the Pre-Intermediate level. Over a four-week period, participants in the experimental group engaged in additional speaking practice through AI chatbots. Changes in oral proficiency were assessed across four dimensions: fluency, grammatical accuracy, lexical resource, and pronunciation. Semi-structured interviews were also conducted to investigate students' experiences of developing speaking skills through chatbot-assisted learning.

The instructional materials were aligned with the Pre-Intermediate proficiency level and adapted for integration with AI-powered learning tools. Task design sought to approximate authentic communicative situations that students might encounter in both everyday and professional contexts. Learning activities included role-play dialogues, situational conversations, discussions of current and professionally relevant topics, and communicative tasks requiring extended oral responses and justification of personal viewpoints.

The following technological tools were employed during the experiment:

AI chatbots: ChatGPT Live, Google Gemini, Duolingo Max, Talkpal AI, Character.AI, Claude AI, and Microsoft Copilot. These platforms were selected for their ability to sustain interactive conversations, provide feedback, and adapt to individual learner needs (Gutiérrez, 2023; Beisembayeva, 2025; Khamsa & Zagidullina, 2025).

Voice-based interfaces: utilized to support pronunciation practice and the development of spontaneous oral communication skills (Gutiérrez, 2023).

Instructional materials: adapted to the Pre-Intermediate level and integrated with chatbot-generated interactive tasks, including role-play scenarios and discussions of everyday topics.

AI chatbots were incorporated into both classroom instruction and independent out-of-class learning activities (Gutiérrez, 2023; Khamsa & Zagidullina, 2025). Students engaged in several forms of interaction:

- **Role-play dialogues**, in which learners practiced communication on assigned topics such as travel, shopping, and social interaction;
- Pronunciation training, using voice-based features to improve articulation and spoken accuracy;
- Spontaneous speaking activities, requiring extended oral responses and open-ended conversations on a variety of topics;
- Individual speaking practice, allowing students to regulate the intensity of interaction and gain additional experience using English in a supportive digital environment.

Particular attention was devoted to reducing communication anxiety and enhancing learners' confidence in spoken English. The opportunity to interact repeatedly with AI chatbots without fear of negative evaluation encouraged greater participation in oral practice and promoted a more active engagement in language learning (Gutiérrez, 2023; Khamsa & Zagidullina, 2025).

Literature review on AI chatbots usage in language learning

Given the growing interest in the integration of artificial intelligence into English language education in the Republic of Kazakhstan, which further highlights the relevance of the present study, a systematic review of international scholarly publications indexed in the Web of Science database was conducted. The list of sources reviewed is provided in the References section.

A comparative analysis of international and Kazakhstani studies suggests that artificial intelligence technologies possess considerable potential for language education. However, their successful implementation depends on a systematic approach that combines technological infrastructure, methodological support, and the development of digital competencies among all participants in the educational process.

Particularly noteworthy is the study conducted by Gutiérrez (2023), which emphasizes that artificial intelligence, and chatbots in particular, creates new opportunities for establishing flexible, safe, and personalized language-learning environments. Tools such as ChatGPT and natural language processing (NLP) technologies enable learners to engage in real-time communication practice, thereby facilitating the development of communicative competence and increasing learner engagement. An especially important finding is that chatbot-assisted interaction may reduce speaking anxiety and provide a supportive environment for independent and out-of-class language practice. At the same time, Gutiérrez identifies several limitations. Most existing chatbots are capable of processing only relatively simple and predictable user requests, reflecting the developmental stage of current language models. Furthermore, concerns remain regarding information accuracy, cultural neutrality, and the ethical implications of AI use in education. Consequently, the author highlights digital literacy among both teachers and learners as a prerequisite for the effective and responsible integration of AI technologies into educational settings (Gutiérrez, 2023).

Another influential contribution is the study "Unleashing the Potential of Chatbots in Education: A State-of-the-Art Analysis" by Winkler and Söllner (2018). This work is of particular theoretical and methodological significance due to its comprehensive review of 1,405 publications addressing the use of chatbots in education, of which 80 studies were selected for detailed analysis. The authors argue that chatbot-supported learning remains at an early stage of development but demonstrates substantial potential for improving learning outcomes and personalizing instruction, particularly within large-scale educational environments. Their study systematizes three core concepts: chatbots, Technology-Mediated Learning (TML), and learning outcomes (Gutiérrez, 2023).

According to Winkler and Söllner, chatbots can be defined as software applications that simulate human conversation through text- or voice-based interaction. Examples include systems such as Alexa, Google Home, and ChatGPT. Chatbots may be classified according to their underlying architecture (retrieval-based or generative), mode of interaction (text or speech), and ability to process conversational context. Chatbot-Mediated Learning (CML) is presented as a specific form of Technology-Mediated Learning characterized by a high degree of personalization, flexibility, real-time interaction, and accessibility (Gutiérrez, 2023).

The authors further categorize learning outcomes associated with chatbot use into three domains: cognitive outcomes (knowledge acquisition and higher-order thinking), affective outcomes (motivation, emotions, and interest), and psychomotor outcomes (practical skills). They also emphasize the contribution of chatbots to the development of metacognitive abilities, particularly through the provision of feedback and support for self-regulated learning. The proposed TML/CML framework consists of three interconnected components: input factors (technology and instructional structure), interaction processes, and expected educational outcomes (Gutiérrez, 2023).

The authors also identify several factors that influence the effectiveness of chatbot-assisted learning:

- Individual characteristics of students (attitudes toward technology, level of self-efficacy, emotional state);
- Educational context (medicine, language learning, massive open online courses);
- Design and functional features of chatbots (flow-based scenarios, artificial intelligence, speech

recognition, contextual data);

- Degree of integration into learning environments (LMS, messaging platforms).

As a result of the analysis, the authors conclude that chatbots can significantly enhance cognitive, affective, and metacognitive learning outcomes, provided that they are implemented appropriately. They are particularly effective in large learning groups where individualized instruction is difficult to ensure, as well as in situations requiring autonomous or flexible modes of interaction.

A more localized perspective is provided by the study of Beisembayeva A.B. (2025), which focuses on the use of artificial intelligence in English language teaching in the Kazakhstani context. The article presents the results of a survey of teachers and students, expert interviews with methodologists and IT specialists, as well as an analysis of the functionality of platforms such as ChatGPT, Grammarly, ELSA Speak, BilimLand, and Daryn.Online. Of particular interest is an experiment conducted at a humanities college in the city of Ust-Kamenogorsk, where the use of AI tools in a learning group led to an 18% increase in vocabulary and grammar proficiency and a 20% reduction in oral errors. Students reported increased motivation and more effective independent learning, while teachers emphasized improved progress tracking and more targeted support. At the same time, the author highlights several limitations, including a lack of technical resources in some regions and the absence of adapted methodological materials. The recommendations include the development of digital infrastructure, professional development for teachers, and support for local AI-enhanced educational platforms (Beisembayeva, 2025).

In the study by Khamsa M.A. and Zagidullina A.A. (2025), titled “Integration of Chatbots in English-speaking skills Education”, published in the journal *News of Ablai Khan Kazakh University of International Relations and World Languages*, the effectiveness of integrating chatbots, particularly ChatGPT and Duolingo Super, into English-speaking skills teaching in higher education institutions of the Republic of Kazakhstan is examined. A combination of theoretical analysis, questionnaire-based research (n=134 teachers and n=237 students), and experimental implementation with B1-level students allowed the authors to identify positive dynamics in the development of language competencies. According to the findings, 82% of teachers and 78% of students reported increased motivation when using chatbots. The experimental group demonstrated significant improvement in both speaking and writing skills, as well as higher engagement levels. The authors emphasize the potential of chatbots for personalizing learning, fostering learner autonomy, and improving the quality of language instruction, while also noting the need for further research to address methodological and technological limitations (Khamsa & Zagidullina, 2025).

Results of the Empirical Study and Discussion

Quantitative Phase of the Study

The experiment was designed to examine changes in students’ oral proficiency and communicative anxiety following regular interaction with AI-based conversational partners and to compare these changes with those observed under traditional speaking instruction. The use of conversational agents in structured learning activities can support academically productive interaction and provide learners with opportunities for sustained dialogue (Tegos et al., 2016).

To test the research hypothesis, the quantitative phase examined three aspects relevant to the expected effects of the intervention: oral proficiency, communicative anxiety, and students’ engagement in speaking activities. Oral proficiency was assessed through speaking fluency, grammatical accuracy, pronunciation, and vocabulary use. Communicative anxiety was measured using the Foreign Language Classroom Anxiety Scale (FLCAS). Students’ engagement in speaking practice was additionally considered through their participation in speaking activities and their responses to the post-intervention questionnaire. The results obtained before and after the intervention were analyzed to determine whether students who participated in AI chatbot-assisted speaking practice demonstrated more pronounced positive changes than those who followed traditional speaking instruction.

An experimental study was conducted involving first-year students enrolled in engineering programmes at Abylkas Saginov Karaganda Technical University. The total sample comprised 80 participants, who were divided into two groups:

Experimental group (n = 40): students who engaged in regular speaking practice using AI chatbots (including ChatGPT Live, Google Gemini, and voice-enabled modules);

Control group (n = 40): students who followed a traditional instruction model without the use of AI-based tools.

Experimental Procedure.

The experimental design consisted of several stages aimed at identifying the impact of AI chatbot-assisted learning on the development of Pre-Intermediate speaking skills:

Pre-test stage (baseline assessment). At the beginning of the experiment, both groups completed a diagnostic assessment that included: assessment of speaking performance in terms of fluency, grammatical accuracy, pronunciation, and interactive participation in dialogue; measurement of communicative anxiety using an adapted version of the Foreign Language Classroom Anxiety Scale (FLCAS).

Four-week intervention phase. The experimental group engaged in AI chatbot-assisted speaking practice three times per week, with each session lasting 30–40 minutes. Activities included role-play dialogues, voice-based exercises, and tasks designed to develop spontaneous speech production.

The control group completed equivalent topic-based speaking activities using traditional instructional methods, including pair work, textbook-based tasks, and classroom discussions.

All speaking activities in both groups were conducted under the supervision of the instructor to maintain comparable instructional conditions.

Post-test stage. After the four-week intervention, both groups completed a post-test identical in structure to the pre-test. This procedure made it possible to compare changes in speaking performance and communicative anxiety over the course of the experiment.

Assessment Instruments

The Foreign Language Classroom Anxiety Scale (FLCAS) (Horwitz et al., 1986) was used to measure communicative anxiety (adapted Russian version);

Students' speaking performance was evaluated according to CEFR-based criteria, including fluency, grammatical accuracy, pronunciation, vocabulary use, and interactive participation in communication;

The quantitative analysis focused on comparing pre-test and post-test results in the experimental and control groups. The obtained data were used to determine the direction and magnitude of changes in the main indicators of speaking performance and communicative anxiety.

Quantitative Results. Table 1 presents the comparative results obtained before and after the four-week intervention.

Table 1.

Comparative results before and after the experiment

Group	Indicator	Pre-test	Post-test	Change (%)
Experimental	Speaking fluency (words per minute)	78	112	+43,6%
Experimental	Accuracy (%)	64%	83%	+19%
Experimental	Communication anxiety (FLCAS)	3,2	2,1	-34,3%
Control	Speaking fluency (words per minute)	76	82	+7,9%
Control	Accuracy (%)	65%	69%	+6,2%
Control	Communication anxiety (FLCAS)	3,3	3,1	-6,1%

The quantitative results demonstrate a more pronounced positive change in the experimental group across the measured indicators. Speaking fluency in the experimental group increased from 78 to 112 words per minute, corresponding to a 43.6% increase. In comparison, the control group demonstrated an increase from 76 to 82 words per minute, or 7.9%.

A similar pattern was observed for grammatical accuracy. In the experimental group, accuracy increased from 64% to 83%, representing a 19-percentage-point improvement. In the control group, accuracy increased from 65% to 69%, corresponding to a 4-percentage-point improvement. Thus, the experimental group demonstrated greater improvement in the measured dimensions of oral proficiency.

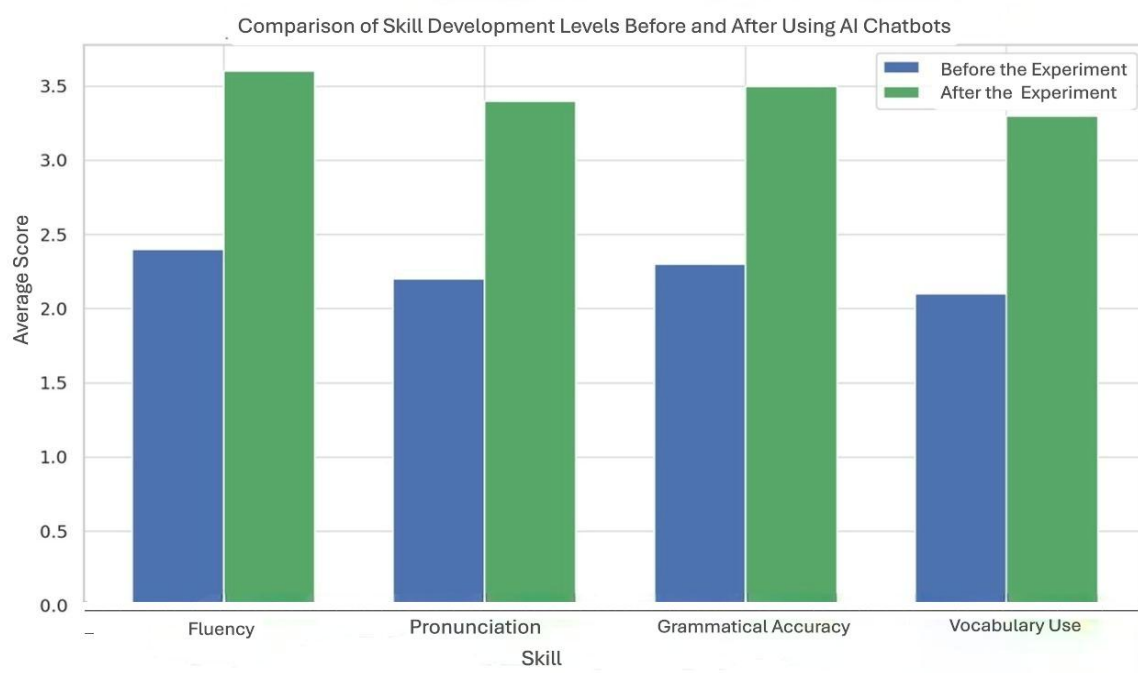
The results for communicative anxiety also showed a more pronounced change in the experimental group. The mean FLCAS score decreased from 3.2 to 2.1, representing a 34.3% reduction. In the control group, the score decreased from 3.3 to 3.1, corresponding to a 6.1% reduction. This difference indicates a substantially greater decrease in communicative anxiety among students who participated in AI chatbot-assisted speaking

practice.

In addition to the comparative measures presented in Table 1, the development of specific speaking skills in the experimental group was assessed using average performance scores. As shown in Figure 1, all four assessed dimensions as fluency, pronunciation, grammatical accuracy, and vocabulary use demonstrated higher average scores after the intervention. The average score for fluency increased from 2.4 to 3.6, pronunciation from 2.2 to 3.4, grammatical accuracy from 2.3 to 3.5, and vocabulary use from 2.1 to 3.3. These results provide additional evidence of positive changes in students' oral proficiency following the four-week intervention.

Figure 1.

Changes in average speaking-skill scores in the experimental group before and after the intervention.



The questionnaire administered after the intervention provided additional information about students' engagement with AI-supported speaking practice. A total of 85% of respondents reported increased confidence in speaking English, 78% indicated perceived improvement in pronunciation, and 90% expressed willingness to continue using AI chatbots for independent English-language learning. These responses indicate a positive perception of the intervention and suggest that students were willing to continue engaging in AI-supported speaking activities beyond the structured classroom tasks.

Taken together, the quantitative findings provide empirical support for the central elements of the research hypothesis. First, the experimental group demonstrated greater improvement in oral proficiency, as reflected in speaking fluency and grammatical accuracy in Table 1 and in the average scores for fluency, pronunciation, grammatical accuracy, and vocabulary use presented in Figure 1. Second, the experimental group demonstrated a substantially greater reduction in communicative anxiety than the control group, as indicated by the FLCAS results. Third, the questionnaire findings showed increased confidence and willingness to continue participating in AI-supported speaking practice, providing additional evidence of students' engagement with the intervention.

Overall, the comparison of pre-test and post-test results indicates that the experimental group experienced more pronounced positive changes than the control group across the principal measured outcomes. These findings provide empirical support for the hypothesis that the use of AI chatbots in speaking instruction can contribute to the development of oral proficiency and the reduction of communicative anxiety among Pre-Intermediate engineering students. The qualitative phase of the study further examined students' experiences and perceptions in order to provide a deeper interpretation of these quantitative findings, including the ways in which AI-supported practice influenced their willingness to communicate and engage in English-speaking activities.

Qualitative Phase of the Study

To explore students' experiences of using chatbots for the development of English-speaking skills, a semi-structured interview was conducted with participants from the experimental group as part of the qualitative phase of the study. While the quantitative data allowed for measuring changes in speaking performance and communicative anxiety levels, the qualitative analysis was aimed at examining students' perceptions, experiences, and attitudes toward the use of AI chatbots, as well as their evaluation of the effectiveness of this form of language practice.

The interviews involved students from the experimental group who had used AI chatbots over a four-week period. The interviews were conducted after the completion of the experiment and included questions related to the usability of AI tools, the impact of digital conversational partners on speaking confidence, features of independent speaking practice, as well as the advantages and limitations of using chatbots in English language learning.

The analysis of the obtained data revealed several recurring thematic categories. The most frequently reported aspect was the opportunity for regular speaking practice without the fear of making mistakes, which contributed to a reduction in psychological barriers when using English. Many participants emphasized that interaction with AI chatbots created a comfortable environment for practicing speaking skills and allowed them to engage in language use at their own pace.

This finding is supported by participants' statements. For example, one respondent noted:

"During communication with the chatbot, I felt more confident because I was not afraid of making mistakes or receiving negative evaluation. This helped me speak more often and more freely." (Respondent 5).

A similar view was expressed by another participant:

"It used to be difficult for me to quickly formulate my thoughts in English, but after several weeks of interaction with the chatbot, it became much easier to respond." (Respondent 11).

A considerable number of respondents also highlighted the positive impact of chatbots on speaking fluency and vocabulary development. According to the students, continuous participation in dialogues facilitated faster retrieval of linguistic resources, improved spontaneous speech skills, and increased confidence in producing extended utterances. The usefulness of voice-based features was also emphasized, as they contributed to the improvement of pronunciation and enabled more natural oral interaction.

In addition, many participants stressed the value of independent practice outside the classroom. Unlike traditional learning formats, AI chatbots provided continuous access to speaking practice and allowed students to regulate the intensity of their learning independently (Oxford, 2017). As one respondent stated:

"I liked that I could study at any time and choose topics that were interesting to me" (Respondent 8).

Another student noted a high level of engagement in the interaction process: *"Sometimes I talked to the chatbot longer than required by the task because it felt like a real conversation rather than a learning exercise"* (Respondent 14).

At the same time, participants identified several limitations in the use of AI tools. These included occasional misinterpretation of prompts, limited depth of some dialogues, and the need to verify certain language suggestions. However, most students considered these drawbacks to be minor compared to the benefits of constant access to speaking practice and immediate feedback.

Overall, the results of the qualitative phase confirmed the findings obtained in the quantitative analysis. The interviews demonstrated that the observed improvement in speaking performance was accompanied by increased language confidence, reduced communicative anxiety, and higher motivation for independent English language learning. The findings suggest that AI chatbots can serve not only as a supplementary tool for language practice but also as an effective means of supporting speaking activity among engineering students.

Integration of Quantitative and Qualitative Results

The use of a mixed-methods research design enabled a multifaceted examination of the impact of AI chatbots on the development of speaking skills. Quantitative data provided evidence of measurable changes in students' oral proficiency, whereas the qualitative analysis helped to explain how learners perceived this experience and which factors, in their view, contributed to the observed outcomes. The comparison of findings

from both phases of the study demonstrated that the results are not contradictory but rather mutually reinforcing and complementary.

The results of the quantitative analysis revealed a consistent positive trend in the experimental group. After four weeks of using AI chatbots, students demonstrated higher levels of speaking fluency and grammatical accuracy, as well as a reduction in communicative anxiety compared with the control group. These findings suggest that regular practice with digital conversational partners may create additional opportunities for the development of English-speaking skills. They are also consistent with previous research on the effectiveness of AI-supported language learning environments (Godwin-Jones, 2019).

During the interviews, we sought to understand the factors underlying these quantitative changes. The analysis of responses showed that most students attributed their progress primarily to the opportunity to use English regularly in a comfortable and psychologically safe environment. Many participants noted that interaction with chatbots made them feel more confident, as they did not fear criticism or negative evaluation. As a result, they engaged more frequently in dialogue, were more willing to use new linguistic structures, and devoted more time to speaking practice.

The relationship between reduced communicative anxiety and increased language confidence appeared particularly significant. While the quantitative results indicated a decrease in FLCAS scores, the interviews provided insight into how these changes manifested in practice. Students repeatedly emphasized that it became easier for them to express their thoughts in English, maintain conversations, and respond to questions without long pauses. Thus, the qualitative data confirmed the trends identified in the quantitative analysis and offered a deeper understanding of students' perceptions of their interaction with AI chatbots.

At the same time, the qualitative analysis revealed certain limitations of AI chatbot use. Some students reported occasional misinterpretations of prompts and a limited depth of discussion on specific topics compared to real-life communication. A number of respondents also highlighted the need for critical verification of some linguistic suggestions, as chatbot responses did not always fully account for communicative context. However, these limitations were sporadic in nature and, according to the majority of participants, did not significantly affect the overall effectiveness of this tool in developing speaking skills.

The integration of quantitative and qualitative data demonstrated the consistency of findings obtained at different stages of the study. The observed improvements in speaking performance were accompanied by positive changes in students' perceptions of their language learning experience. Previous research has also highlighted the didactic potential of chatbots as supplementary tools in English language teaching (Lobeeva, 2023). In the context of the present study, these findings support the view that AI chatbots can provide additional opportunities for speaking practice and contribute to a more flexible learning environment for engineering students. These findings allow for a more detailed discussion of the pedagogical implications of AI chatbot use and the potential mechanisms underlying their impact on the development of speaking skills among engineering students.

Conclusion

The findings of the present study provide empirical support for the hypothesis that the use of AI chatbots can contribute to the development of speaking skills among technical university students at the Pre-Intermediate level. Based on theoretical analysis, a review of contemporary applied linguistics literature, and empirical data obtained during the pedagogical experiment, it was established that regular interaction with AI conversational partners was associated with improvements in key indicators of speaking performance, including fluency and grammatical accuracy, as well as a reduction in communicative anxiety.

The experiment conducted at Abylkas Saginov Karaganda Technical University involved 80 first-year engineering students who were divided into experimental and control groups. Over a four-week period, students in the experimental group interacted with AI chatbots (ChatGPT Live, Google Gemini, etc.) within structured speaking tasks aimed at developing both dialogic and monologic speech, including voice-based exercises and role-play activities. All interactions between students and chatbots were conducted during in-class time under the supervision of the instructor.

The quantitative analysis indicated positive changes in the experimental group. Speaking fluency increased from 78 to 112 words per minute, while grammatical accuracy improved from 64% to 83%. Communicative anxiety, measured using the FLCAS, decreased from 3.2 to 2.1. In comparison, the control

group demonstrated more limited changes across all three indicators.

The quantitative findings were complemented by the results of the qualitative phase of the study, which made it possible to identify key mechanisms underlying the observed changes. Students reported that interaction with AI chatbots created a psychologically safe environment for speaking practice, reduced their fear of making mistakes, and contributed to more confident communicative behavior. The availability of immediate feedback and regular opportunities for speaking practice were also identified as important factors.

The participants also associated improvements in their speaking skills not only with increased practice frequency but also with a more positive attitude toward English communication. Many students demonstrated greater readiness to communicate, reduced tension during spontaneous speech, and increased motivation to use English in both academic and everyday contexts.

At the same time, the qualitative analysis revealed certain limitations of AI chatbot use, including occasional misinterpretation of prompts, limited depth of some interactions, and the need for critical evaluation of linguistic feedback provided by the system. These limitations were generally perceived by the participants as manageable within the learning context and did not prevent them from using AI chatbots as an additional tool for speaking practice.

Overall, the findings suggest that the use of AI chatbots can provide additional opportunities for speaking practice and contribute to a more flexible, personalized, and psychologically comfortable learning environment conducive to the development of communicative competence. The findings also indicate that AI technologies may influence both cognitive and affective aspects of English language learning. AI chatbots therefore have potential as a supplementary tool in the language training of engineering students. At the same time, their implementation requires further methodological refinement, expansion of the experimental base, and investigation of long-term effects across different proficiency levels.

Conflict of Interest Statement

Authors declare no potential conflicts of interest with respect to the research, authorship, or publication of this article.

Author Contributions

Damira Jantassova and Eleonora Prozovskaya contributed equally to the conceptualization, methodology, data collection and analysis, literature review, preparation of the original manuscript draft, visualization, validation, review, editing, and finalization of the study. Both authors participated in all stages of the research and approved the final version of the manuscript.

Declaration of Use of Generative AI and AI-Assisted Technologies

The authors declare that no generative AI or AI-assisted technologies were used in the preparation of this manuscript.

References

- Beisembayeva, A. B. (2025). Artificial intelligence in English language teaching in Kazakhstan: Opportunities and prospects. *Bulletin of Zhetysu University*, 1(114), 60-65. <https://doi.org/10.53355/ZHU.2025.114.1.007>
- Character.AI. (2022). Character.AI chat platform. <https://character.ai/>
- Godwin-Jones, R. (2019). Emerging technologies: Language learning and artificial intelligence. *Language Learning & Technology*, 23(3), 8-12. <https://doi.org/10.1017/S0958344019000104>
- Google Cloud. (2023). Dialogflow: Build conversational interfaces for websites, mobile applications, and messaging platforms. <https://cloud.google.com/dialogflow>
- Gutiérrez, L. (2023). Artificial intelligence in language education: Navigating the potential and challenges of chatbots and NLP. *Research Studies in English Language Teaching and Learning*, 1(3), 180-191. <https://doi.org/10.62583/rselt.v1i3.44>
- Khamsa, M. A., & Zagidullina, A. A. (2025). Integration of chatbots in English-speaking skills education. *News of Ablai Khan Kazakh University of International Relations and World Languages. Series: Pedagogical Sciences*, 1(76), 42-57. <https://doi.org/10.48371/PEDS.2025.76.1.003>

- Lobeeva, P. I. (2023). The didactic potential of using chatbots in teaching and learning English phrasal verbs. *Tambov University Review. Series: Humanities*, 28(6), 1467-1476. <https://doi.org/10.20310/1810-0201-2023-28-6-1467-1476>
- OpenAI. (2024). ChatGPT for language learners. <https://openai.com/chatgpt>
- Oxford, R. L. (2017). *Teaching and researching language learning strategies*. Routledge.
- Talkpal AI. (2023). Talkpal AI language learning platform. <https://talkpal.ai>
- Tegos, S., Demetriadis, S., Papadopoulos, P. M., & Weinberger, A. (2016). Conversational agents for academically productive talk: A comparison of directed and undirected agent interventions. *International Journal of Computer-Supported Collaborative Learning*, 11(4), 417-440. <https://doi.org/10.1007/s11412-016-9246-2>
- Winkler, R., & Söllner, M. (2018). Unleashing the potential of chatbots in education: A state-of-the-art analysis. *Academy of Management Proceedings*.

Information about authors

Damira Jantassova - Candidate of Pedagogical Sciences, Professor, Head of the Department of Foreign Languages, Abylkas Saginov Karaganda Technical University, Karaganda, 100027, Republic of Kazakhstan; e-mail: d.dzhantasova@ktu.edu.kz; ORCID iD: 0000-0003-2595-3249 (*corresponding author*).

Eleonora Prozovskaya - Master of Foreign Philology, Senior Lecturer, Department of Foreign Languages, Abylkas Saginov Karaganda Technical University, Karaganda, 100027, Republic of Kazakhstan; e-mail: e.prozovskaya@ktu.edu.kz; ORCID iD: 0009-0008-1345-4307.

Received: 11.06.2026

Revised: 07.09.2026

Accepted: 29.09.2026

Published: 30.09.2026

Type of article: Research Article

¹Askat Tleuov*, ²Mukhametzhan Seitzhapparuly, ³Kairat Moldashev

¹ KIMEP University, College of Human Sciences and Education, Republic of Kazakhstan

² Narxoz University, Republic of Kazakhstan

³ SDU University, Republic of Kazakhstan

UNEVEN PARTNERSHIPS: PHD STUDENTS' SATISFACTION WITH LOCAL AND FOREIGN SUPERVISORY FEEDBACK AND INTERNATIONALIZATION OUTCOMES IN KAZAKHSTAN

Abstract. Kazakhstan's doctoral education policy requires all PhD students to have both a local and a foreign supervisor. This co-supervision model was designed to connect doctoral training with international academic standards. Little empirical work has asked what this arrangement means for students' experience of supervision, or whether satisfaction with feedback from each supervisor type tracks internationalization outcomes. Survey data were collected from 845 doctoral students and graduates across institutions in Kazakhstan. Valid sample sizes varied across analyses, ranging from 586 to 828 cases, because the satisfaction items included response options such as 'not applicable' and 'prefer not to answer'. Independent samples t-tests, chi-square tests, and Spearman rank-order correlations were used to test associations between supervisor feedback satisfaction and three outcomes: Scopus publication output, international conference participation, and international internship outcomes. Satisfaction with foreign supervisor feedback predicted higher Scopus output ($t = 3.06$, $p = .002$) and conference participation ($\chi^2 = 3.85$, $p = .049$), and correlated moderately with internship outcomes ($\rho = .56$). For local supervisor feedback, no comparable association appeared with publication output or conference attendance; only the internship correlation reached significance, and at a notably weaker magnitude ($\rho = .37$). Internationalization in Kazakhstan's PhD system, these results indicate, runs mainly through the foreign half of the co-supervision arrangement. The local half contributes far less to the outcomes the policy was designed to produce from both sides. What the model needs is not a new mandate but the institutional conditions, reduced teaching loads and dedicated mobility funding chief among them, that would let local supervisors operate on comparable terms.

Keywords: internationalization of higher education; doctoral supervision; co-supervision; supervisor feedback satisfaction; doctoral internationalization outcomes; Kazakhstan.

Introduction

Kazakhstan joined the Bologna Process in 2010, and doctoral education has since been positioned as a primary site of the country's internationalization agenda. The term internationalization carries different meanings and priorities across national systems (Altbach & Knight, 2007); in Kazakhstan, the emphasis has been on connecting domestic doctoral training to global knowledge production and academic networks (Sagintayeva & Kurakbayev, 2015; Kuzhabekova, Bilyalov, et al., 2022).

The connection between doctoral training and global academic networks runs primarily through supervisors. Lee (2008) shows that supervisory input extends well beyond dissertation assessment: it shapes students' publication decisions, conference access, and socialization into research communities. In Kazakhstan's reform model, that connection is the co-supervision arrangement's primary purpose.

Kazakhstan's response was structural: every doctoral candidate must have both a local and a foreign supervisor. The foreign supervisor is formally responsible for connecting students to international publishing norms and academic networks; the local supervisor provides orientation to national academic conventions. Those two responsibilities are not equally resourced. Local supervisors typically carry heavier teaching loads, receive less institutional support, and maintain weaker connections to international research networks than their foreign counterparts (Sagintayeva & Kurakbayev, 2015). The policy's complementary design therefore depends on conditions it does not create.

Students' experiences of co-supervision reflect this asymmetry. Satisfaction with doctoral supervision

is shaped by three conditions: clarity about supervisory responsibilities, quality and frequency of interaction, and shared understanding of academic norms (Lee & Green, 2009; Manathunga, 2014). The co-supervision arrangement does not guarantee a balanced partnership, and the unevenness it can produce carries a theoretical stake beyond individual student experience: if internationalization outcomes depend mainly on the foreign supervisor, the model may generate outputs without building the local capacity needed to sustain them independently, a possibility the Literature Review develops through Shahjahan and Kezar's (2013) account of dependency in international academic partnerships.

That possibility is examined here using survey data from doctoral students across Kazakhstani institutions. The study investigates associations between PhD students' satisfaction with feedback from local and foreign supervisors and three internationalization outcomes: Scopus publication output, international conference participation, and international internship outcomes. The central research question is: "How is satisfaction with local and foreign supervisor feedback associated with doctoral students' internationalization outcomes in Kazakhstan?"

Gaps and Contribution

Most research on internationalization in Kazakhstan has examined the faculty perspective (Kuzhabekova et al., 2019; Kuzhabekova, Bilyalov, et al., 2022) or analyzed policy reform at the institutional level (Sagintayeva & Kurakbayev, 2015). How doctoral students assess the feedback they receive from each supervisor type, and whether that satisfaction relates to measurable internationalization outcomes, has received far less attention. This study addresses that gap by examining associations between PhD students' satisfaction with local and foreign supervisor feedback and three outcomes: Scopus publication output, international conference participation, and international internship outcomes.

The contribution extends beyond description. By disaggregating satisfaction by supervisor type, the study makes it possible to test whether internationalization outcomes are distributed across both sides of the co-supervision partnership or concentrate on one. If foreign supervisor feedback satisfaction consistently predicts outcomes that local supervisor feedback does not, the data provide empirical grounding for the argument that Kazakhstan's co-supervision model reproduces rather than reduces the hierarchies it was designed to dismantle.

Contextual Background. Doctoral Education in Kazakhstan

For decades, doctoral training in Kazakhstan followed the Soviet-inherited candidate of sciences model. Advancement within that model served the reproduction of national academic institutions more than entry into international circuits of knowledge production. Research output did not normally require indexed international publication, and graduates were assessed against nationally defined criteria with limited currency outside the post-Soviet academic space (Sagintayeva & Kurakbayev, 2015). The move to a three-cycle, Bologna-compatible PhD changed the terms of doctoral completion. Activities that had previously been peripheral, including indexed journal publication, conference participation, and research mobility abroad, became formal requirements. These requirements were linked to Kazakhstan's policy ambition to position research universities as contributors to a knowledge economy (Kuzhabekova et al., 2022).

The system that received these requirements had not been built for them. Soviet-era universities had separated teaching from research, concentrated research funding in a small group of flagship institutions, and maintained weak ties to international publishing and conference networks (Sagintayeva & Kurakbayev, 2015). Domestic incentive structures had rewarded national output more consistently than articles in indexed international journals (Kuzhabekova et al., 2019). The publication requirement therefore arrived before many universities had developed the institutional capacity to support it. Kuzhabekova (2025) shows the effect clearly: Kazakhstan's PhD publication requirement has increased the number of articles produced by doctoral students, but their average research impact has declined. This pattern points to publication under pressure, not to a corresponding expansion of research capacity.

Students describe the publication rule as difficult to meet without institutional backing. They report limited access to international databases, uneven English-language support, and supervisory guidance that does not always extend to international journal submission (Yelibay et al., 2022). In these accounts, the problem shifts from the requirement itself to the institutional scaffolding needed to make the requirement achievable through local resources.

The practical operation of co-supervision has begun to receive more direct empirical attention.

Tokatligil et al. (2025), studying the arrangement from the perspectives of both Kazakhstani and international co-supervisors, describe it as an intercultural contact zone where collaboration and coordination problems coexist. Kalman et al. (2022), in their review of co-supervision across doctoral systems, identify internationalization as one of four institutional drivers behind its spread. Their review also shows that coordination problems and role ambiguity become more acute when supervisors work from different national academic systems. Kazakhstan's mandatory model sits at this intersection: internationalization supplies the official rationale, and supervisory asymmetry becomes a foreseeable structural consequence.

Supervisors and Their Role in Doctoral Internationalization

Doctoral supervision needs to offer more than broad support if it is to achieve the outcomes envisioned in Kazakhstan's reform agenda. Halse and Malfroy (2010) describe supervision as a form of professional work consisting of five elements: a learning alliance, habits of mind, scholarly expertise, technical skill, and contextual expertise. The final element is especially important for the internationalization of doctoral education. It involves understanding which journals are influential, which conferences are valuable, how peer review operates, and which professional networks can help doctoral students become part of a wider research community. A supervisor may be strong pedagogically, intellectually, and technically, yet still lack this particular kind of international research knowledge.

This distinction helps clarify the imbalance between local and foreign supervisors. Many local supervisors have had limited opportunities or incentives to develop expertise in international journal submission and peer review, because the academic systems in which they advanced did not require such experience (Kuzhabekova et al., 2019). The issue is therefore institutional rather than personal. It reflects a form of contextual expertise that domestic academic careers in Kazakhstan have not traditionally valued or rewarded.

Within this arrangement, foreign supervisors are expected to bring knowledge of international publishing standards, conference networks, and peer-review practices. This does not mean that they are generally better supervisors. Rather, they are often better placed to provide the specific contextual knowledge on which the reform's international ambitions depend.

This unevenness is not experienced by students alone. Tokatligil et al. (2025) show that international co-supervisors often feel uncertain about the boundaries of their role and receive limited guidance about what their Kazakhstani colleagues expect from them. These conditions affect the consistency of their engagement with the doctoral students assigned to them. As a result, a model based on complementary supervision may leave both local and international supervisors operating with only a partial understanding of what the arrangement actually requires.

Foreign Internship and Publication Requirements in Kazakhstan's Doctoral Framework

Kazakhstan's doctoral framework translates internationalization into three formal markers of completion: publication in Scopus-indexed journals, participation in international conferences, and an international research internship or extended stay abroad. The publication rule carries the strongest institutional force because it is tied directly to degree conferral. Since 2011, regulations have required doctoral candidates to publish at least two articles in journals indexed in Web of Science or Scopus, with the requirement becoming progressively stricter over time (Kuzhabekova, 2025). Many universities have also introduced higher minimums in their own institutional regulations, making indexed publication a condition of graduation rather than a general aspiration (Kuzhabekova, Bilyalov, et al., 2022).

In practice, doctoral internationalization in Kazakhstan is not an optional enrichment of doctoral study. It is built into the rules of completion. It is measured through articles, conference visibility, and time spent in an international research setting. Every doctoral candidate must meet these markers, but whether the supervisory resources needed to do so are distributed evenly across local and foreign supervisory roles is the question this study takes up.

Literature Review

Kazakhstan's regulations give internationalization a precise institutional form, but the concept itself carries no such settled definition across higher education systems more broadly. Altbach and Knight (2007) describe it as the incorporation of international, intercultural, and global elements into higher education, but the motivations vary: for some institutions, the goal is educational quality, for others it is reputation and market positioning, and for others it is capacity building. A capacity-building rationale evaluates a policy by whether

domestic research capacity grows; a marketization rationale evaluates it by institutional visibility and rankings. These distinctions are not interchangeable, and which motivation drives a given reform shapes what counts as evidence that it has succeeded.

Doctoral education is a primary site of this debate. PhD training functions as both an apprenticeship in research and a mechanism through which the publication standards, methodological conventions, and professional norms of international science are transmitted and reproduced (Kehm & Teichler, 2007). At the doctoral level, joint programs, student mobility, and co-supervision are the main internationalization mechanisms. All three treat best practice as flowing from established to emerging systems; local traditions of doctoral training become deficits to correct rather than legitimate alternatives (Nerad & Evans, 2014). Such reforms produce convergence without building the local expertise to sustain it. External expertise does not function as a transitional scaffold; the model has no mechanism for making it dispensable.

Supervisory relationships are where this directional logic operates most concretely. Three outcomes recur in this literature as markers of doctoral internationalization : indexed journal publication, international conference participation, and research mobility experiences such as internships and extended stays abroad. Lee (2008) shows that supervisors do more than assess dissertations; access to each of these outcomes depends substantially on the supervisor's own research networks and knowledge of international publishing conventions. Evidence from European and Asian universities shows that supervisors with international experience are more likely to direct students toward indexed journal publication and international career trajectories (McAlpine & McKinnon, 2013). Co-supervision does not resolve this unevenness by adding a second supervisor; it displaces the question onto the arrangement itself. Which supervisor delivers which outcomes becomes the operative question.

Shahjahan and Kezar (2013) argue that international academic collaborations frequently reproduce hierarchies in which scholars from less-established systems are positioned as junior partners, dependent on foreign expertise with no structural mechanism for building autonomous research capacity. The argument applies directly to co-supervision: if responsibility for internationalization outcomes falls primarily on the foreign supervisor, the model does not build a sustainable partnership; it creates a structural dependency in which local supervisors are formally present but functionally marginal. Outcomes may still accrue to students, but they do so through the foreign supervisor's networks and expertise, not through any capacity built within the national academic system.

Kazakhstan's co-supervision policy sits directly within this tension. As part of its Bologna Process reforms, Kazakhstan mandated that all doctoral candidates have both a local and a foreign supervisor, assigning the foreign supervisor explicit responsibility for connecting students to international publishing norms and academic networks (Sagintayeva & Kurakbayev, 2015). The policy is premised on the two supervisors working as complementary partners. Faculty members have described internationalization as an externally imposed agenda that does not adequately account for local academic conditions (Kuzhabekova, Bilyalov, et al., 2022). Local supervisors face specific constraints that limit international engagement: heavy teaching loads, limited mobility funding, and weak institutional incentives (Kuzhabekova et al., 2019). Student experiences vary. Some report gaining access to international research networks and publishing guidance through their foreign supervisors; others describe irregular communication, misaligned expectations, and unequal engagement between the two parties (Kuzhabekova et al., 2019). Co-supervision neither consistently delivers on its rationale nor consistently fails; what it produces depends on each supervisor's resources, network connections, and willingness to coordinate.

Students' ratings of their supervisors predict task engagement and research performance (Wu et al., 2025). Those who perceive their supervisor as academically engaged report higher scholarly commitment and productivity. In a co-supervision model where one supervisor has substantially greater access to international research networks than the other, ratings of the two supervisors are unlikely to converge. Differential satisfaction may track which side of the arrangement is actually delivering the publication, conference, and internship outcomes the policy targets.

Supervisor Feedback and Publication Output

Three hypothesis pairs follow from the literature reviewed, each corresponding to one of the doctoral internationalization outcomes examined.

Doctoral publication output is shaped substantially by supervisory input. Students who receive specific

guidance on research design, academic writing, and publication conventions produce more indexed journal articles (Khuram et al., 2023; Wilkins et al., 2021; Shen & Jiang, 2023). The character of that guidance also matters: directive supervision predicts short-term publication productivity; more autonomous approaches develop research capacity over time but typically yield slower output (Wu et al., 2025). Students' satisfaction with supervisor feedback may therefore index the degree to which supervisors are providing the structured, publication-oriented guidance that indexed journal output requires.

H1a: PhD students who are satisfied with foreign supervisor feedback report higher Scopus publication output than those who are not satisfied.

H1b: PhD students who are satisfied with local supervisor feedback report higher Scopus publication output than those who are not satisfied.

International Conference Participation

Conference participation in doctoral education is not self-directed. Supervisors actively shape whether and where students present their work, guiding them toward conferences that fit their research trajectory and professional goals (Haus, 2021). Students supervised by academics with established conference networks are better positioned to receive specific recommendations, gain introductions, and build professional connections that attendance makes possible (Orellana et al., 2016). Satisfaction with supervisor feedback may therefore capture more than an evaluation of dissertation guidance; it may reflect whether students perceive their supervisors as actively connecting them to international conference networks.

H2a: Satisfaction with foreign supervisor feedback is associated with higher rates of international conference participation among PhD students.

H2b: Satisfaction with local supervisor feedback is associated with higher rates of international conference participation among PhD students.

International Internship Outcomes

International internship and research stay experiences depend partly on supervisory preparation; the quality of guidance students receive before and during such placements shapes how they experience those settings (Mainhard et al., 2009). Students who receive structured supervisory support in unfamiliar research environments report stronger professional outcomes and a clearer sense of scholarly identity (Vattøy, 2026). Supervisors with established international connections also help students access internship placements by linking them to research networks abroad. Satisfaction with supervisor feedback may therefore index how well students feel their supervisors have prepared them for international research settings and connected them to the opportunities those settings offer.

H3a: Satisfaction with foreign supervisor feedback is positively correlated with PhD students' satisfaction with foreign internship experiences.

H3b: Satisfaction with local supervisor feedback is positively correlated with PhD students' satisfaction with foreign internship experiences.

Methodology

Data Collection and Ethical Considerations

The analysis draws on survey data collected by the National Academy of Sciences of the Republic of Kazakhstan from doctoral students and doctoral graduates across higher education institutions and research organizations in Kazakhstan. Data were collected online through the SurveyMonkey platform between July 20 and August 11, 2025. The survey link and invitation letters were sent to higher and/or postgraduate education institutions and research organizations in Kazakhstan for further distribution among potential respondents. In addition, invitation letters were sent directly to doctoral students and doctoral graduates using email addresses collected from publications in academic journals or provided by research organizations.

The authors received formal authorization and were granted full access to the dataset for research purposes. The survey was conducted anonymously. No direct personal identifiers were collected, and demographic information was used only for classification and aggregate analysis. The study received ethical approval in July 2025. An informed consent form was presented to all potential respondents on the first page of the SurveyMonkey questionnaire, with two response options: "I agree" and "I do not agree." Respondents who selected "I do not agree" were directed to the end of the questionnaire, while those who selected "I agree" proceeded to complete the survey.

Variables and Measures

The initial dataset contained 845 survey responses. The sample included current doctoral students and recent doctoral graduates. For readability, the term doctoral respondents is used throughout unless a distinction is analytically relevant. The respondents were asked about their satisfaction with feedback from both local and foreign supervisors, their publication record, particularly the number of Scopus-indexed publications, participation in international conferences, and satisfaction with foreign internship experiences.

Satisfaction-related items used the following response scale: 4 = satisfied, 3 = mostly satisfied, 2 = mostly not satisfied, 1 = not satisfied, 96 = not applicable, 97 = did not need, and 98 = prefer not to answer. Responses coded as 96, 97, and 98 were treated as non-substantive responses and excluded from analyses involving the relevant variable. This approach was used because these categories do not represent substantive evaluations of satisfaction.

For analyses requiring a binary satisfaction measure, responses of 4 (“satisfied”) and 3 (“mostly satisfied”) were grouped into the “satisfied” category, while responses of 2 (“mostly not satisfied”) and 1 (“not satisfied”) were grouped into the “not satisfied” category. For analyses using satisfaction as an ordinal variable, the original 1–4 substantive scale was retained, and non-substantive responses were excluded.

Because non-substantive responses were excluded only for the variables involved in each analysis, the number of valid cases varies across the results tables. This means that the study did not apply a single fixed sample size to all analyses; instead, each analysis used variable-specific valid cases.

The number of Scopus-indexed publications was treated as a continuous variable. Participation in international conferences was treated as a binary categorical variable, coded as 1 if the respondent had attended an international conference and 0 if the respondent had not attended.

Statistical Analysis

The analysis was conducted using Python. Prior to analysis, relevant variables were converted into numeric format, and non-substantive response codes were excluded from the specific analyses in which those variables were used.

To examine whether satisfaction with supervisor feedback was associated with publication output, Welch’s independent-samples t-tests were conducted. Welch’s t-test was used because it does not assume equal variances between the comparison groups. To examine whether satisfaction with supervisor feedback was associated with participation in international conferences, chi-square tests of independence were conducted. These tests were appropriate because both variables were categorical: supervisor feedback satisfaction was recoded into satisfied and not satisfied groups, and conference participation was coded as attendance or non-attendance.

Finally, Spearman’s rank-order correlation was used to examine the relationship between satisfaction with supervisor feedback and satisfaction with foreign internship experiences. Spearman’s correlation was selected because the satisfaction variables were ordinal and the analysis did not require assumptions of linearity or normally distributed variables.

Results

Supervisor Feedback and Publication Output

In Table 1, doctoral students who are satisfied with feedback from foreign supervisors ($n = 665$) have a higher mean number of Scopus publications ($M = 1.98$) than those who are not satisfied ($n = 106$; $M = 1.22$). This difference is statistically significant, $t = 3.06$, $p = 0.0024$. The small p -value suggests that this difference is unlikely to be due to sampling variation alone. By contrast, there is no statistically significant difference between those who are satisfied with feedback from a local supervisor ($n = 725$; $M = 1.85$) and those who are not ($n = 103$; $M = 1.67$), $t = 0.50$, $p = 0.6159$ (Table 2).

The total sample size is 845, of which 74 indicated “prefer not to answer” for satisfaction with Foreign Supervisor Feedback and 17 chose this option for satisfaction with Local Supervisor Feedback. This might be due to the fact that some PhD students were not yet assigned a local/foreign supervisor or haven’t formed an opinion. Given that it takes more time to find a foreign supervisor, the higher number of “prefer not to answer” for this group is based on practical reasons.

Table 1.*T-test Results: Scopus Publications vs. Foreign Supervisor Feedback*

Number in 'Satisfied' group: 665 Average Publications for 'Satisfied' group: 1.98	Number in 'Not Satisfied' group: 106 Average Publications for 'Not Satisfied' group: 1.22
T-statistic: 3.0641; P-value: 0.0024	

Table 2.*T-test Results: Scopus Publications vs. Local Supervisor Feedback*

Number in 'Satisfied' group: 725 Average Publications for 'Satisfied' group: 1.85	Number in 'Not Satisfied' group: 103 Average Publications for 'Not Satisfied' group: 1.67
T-statistic: 0.5028; P-value: 0.6159	

International Conference Participation

A statistically significant connection was found between satisfaction with feedback from foreign supervisor and participation in international conferences (Table 3), $\chi^2(1) = 3.85$, $p = .049$. The students who reported satisfaction with foreign supervisor feedback were also more likely to have attended international conferences. Because the effect size was small (Cramer's $V = .07$), the practical significance is small. However, no statistically significant connection between satisfaction with feedback from local supervisor and participation in international conferences (Table 4), $\chi^2(1) = 1.74$, $p = .187$ was found. The magnitude of association, Cramer's $V = .05$, was negligible.

Table 3.*Satisfaction with Foreign Supervisor Feedback and International Conference Participation*

Feedback satisfaction	Did not attend	Attended	Total
Not satisfied	65	41	106
Satisfied	336	329	665
Total	401	370	771
Chi-Square Statistic: 3.8466; P-value: 0.0498; Cramer's $V = .07$			

Table 4.*Satisfaction with Local Supervisor Feedback and International Conference Participation*

Feedback satisfaction	Did not attend	Attended	Total
Not satisfied	61	42	103
Satisfied	375	350	725
Total	436	392	828
Chi-Square Statistic: 1.7448; P-value: 0.1865; Cramer's $V = .05$.			

International Internship Outcomes

To test Hypothesis 3, Spearman's rank-order correlation analyses were conducted to examine the association between PhD students' satisfaction with supervisory feedback and their satisfaction with foreign internship experiences. As indicated in Table 5, there was a significant moderate positive correlation between students' satisfaction levels with feedback from foreign supervisors and their satisfaction levels with foreign internships ($\rho = .56$, $p < .001$). That is, the greater the satisfaction with foreign supervisory feedback, the greater the satisfaction with international internship experiences. There was also a statistically significant but weaker positive relationship between satisfaction levels with local supervisor feedback and satisfaction levels

with foreign internships ($\rho = .37, p < .001$). The magnitude of this correlation is substantially smaller than that between foreign supervisors and internship experiences.

Table 5.

Spearman Correlations Between Supervisor Feedback Satisfaction and Satisfaction with International Internship Outcomes

Supervisor feedback satisfaction	International Internship Outcomes	n	Spearman's ρ	p
Foreign supervisor feedback	Foreign internship	586	0.560	< .001
Local supervisor feedback	Foreign internship	614	0.366	< .001

Discussion

Publication output differentiates clearly by supervisor type. Students satisfied with foreign supervisor feedback reported a higher mean number of Scopus-indexed publications than those who were not satisfied ($M = 1.98$ vs $M = 1.22$), a difference that is statistically significant ($t = 3.06, p = .002$). For local supervisor feedback satisfaction, no corresponding difference emerged ($M = 1.85$ vs $M = 1.67, t = 0.50, p = .616$). The result aligns with what McAlpine and McKinnon (2013) and Lee (2008) suggest: supervisors with established international connections and publishing experience are better positioned to orient students toward indexed journal publication. Whether the association reflects a direct supervisory effect or whether students who publish more tend to evaluate their supervisors more favourably, the cross-sectional design cannot establish. What the finding does show is that publication output, the most quantifiable marker of doctoral internationalization in this study, is differentially distributed across the two sides of Kazakhstan's co-supervision arrangement.

Foreign supervisor feedback satisfaction is also associated with international conference attendance ($\chi^2(1) = 3.85, p = .049$), though the effect is small (Cramer's $V = .07$). Satisfaction with local supervisor feedback shows no comparable association ($\chi^2(1) = 1.74, p = .187$). The effect size limits what this can support: at Cramer's $V = .07$, foreign supervisor feedback satisfaction is a modest contributor to conference attendance, and the data do not permit stronger claims.

The internship findings depart from the pattern observed for the other two outcomes. Both supervisor types show statistically significant correlations with internship outcomes, but at different magnitudes: foreign supervisor feedback satisfaction correlates at $\rho = .56$ ($p < .001$), local supervisor feedback satisfaction at $\rho = .37$ ($p < .001$). This is the only finding in which local supervisor feedback satisfaction shows a statistically significant association with any internationalization outcome. Internship preparation may draw on forms of supervisory support that local supervisors are partially positioned to provide, even without direct access to international research networks. The foreign supervisor correlation remains substantially stronger, which is consistent with the expectation that supervisors with established international connections are better placed to prepare students for research environments abroad. Both associations are real; the gap between them marks the limit of what local supervisors contribute to international mobility experiences compared with their foreign counterparts.

Measured against three internationalization outcomes, Kazakhstan's co-supervision model does not operate as a balanced partnership. Publication output and conference participation show statistically significant associations only with foreign supervisor feedback satisfaction; for internship outcomes, both supervisor types register, but the foreign supervisor correlation is substantially stronger ($\rho = .56$ vs $\rho = .37$). Kazakhstan's policy designates both supervisors as equal partners, but the pattern in the data suggests the partnership is structured unevenly in practice. The internationalization function the model was designed to deliver concentrates on one of its two components.

Attributing this asymmetry to individual supervisory quality would misread the evidence. Local supervisors in Kazakhstan typically face institutional constraints that their foreign counterparts do not: heavier teaching loads, limited mobility funding, and weaker connections to international research networks (Kuzhabekova et al., 2019; Sagintayeva & Kurakbayev, 2015). These are not deficits of supervision; they are conditions under which local supervisors work.

Shahjahan and Kezar (2013) argue that international academic partnerships frequently reproduce

hierarchies in which scholars from less-established systems become dependent on foreign expertise rather than building autonomous research capacity. The findings here are consistent with that argument, though they cannot confirm it directly: the data show associations, not mechanisms. What they make visible is a structure in which doctoral students' access to publication networks, conference circuits, and international research stays runs primarily through the foreign supervisory relationship.

Policy Implications

The findings identify a design problem in Kazakhstan's co-supervision system. The mandate can help current doctoral students reach publication, conference, and mobility targets. It does not, by itself, build the local supervisory capacity that would make those gains durable.

That distinction matters. A policy can produce outputs while preserving the conditions that make external support necessary.

The most immediate issue is institutional. Kazakhstan's doctoral regulations name local and foreign supervisors as equal partners, yet the data do not show equal operation. Publication output and international conference attendance are statistically associated only with satisfaction with foreign supervisor feedback. Internship outcomes are associated with satisfaction with both supervisors, with a stronger relationship for foreign supervisor feedback than for local supervisor feedback ($p = .56$ compared with $p = .37$).

A simple "supervisor quality" explanation would misread the problem. Local supervisors work within conditions that restrict the practices the mandate expects from them. Heavy teaching and administrative loads leave limited time for research, manuscript development, and sustained doctoral feedback, a problem documented in earlier OECD reporting and in subsequent empirical studies of Kazakhstani academics (OECD, 2017; Kuzhabekova & Ruby, 2018; Moldashev & Tleuov, 2022). Faculty also report weaker access to international mobility and research collaboration opportunities than students, which limits the professional networks through which conference participation, publication collaboration, and external supervisory ties are usually built (Kuzhabekova, Ispambetova, et al., 2022; Kuzhabekova, Bilyalov, et al., 2022). Equal partnership cannot be produced by regulation alone. It requires workload, funding, and evaluation structures that give local supervisors the time, networks, and career reasons to perform the internationalizing work assigned to them.

Reducing teaching loads and funding mobility for local doctoral supervisors should be treated as operating conditions for the policy, not optional additions. Manuscript feedback is slow work. It involves choosing journals, revising arguments, responding to reviewers, interpreting editorial decisions, and helping students recover from rejection. Conference mentoring also requires more than approving an abstract. Supervisors need access to the meetings where research networks, future collaborations, and publication opportunities take shape. A local supervisor carrying a full teaching load cannot reliably provide that level of iterative support. A local supervisor without conference funding cannot easily build the international relationships that doctoral students are later expected to enter (Wilkins et al., 2021).

Under these conditions, the foreign supervisor's role shifts from complement to compensation. The foreign supervisor supplies the international publishing and networking resources that the local system has not yet made available to its own supervisors. That arrangement can sustain student outputs. It also leaves the underlying imbalance intact. Reducing teaching loads and creating mobility funding for local supervisors would begin to move the model toward the complementary partnership described in policy rather than the compensatory arrangement visible in practice.

The same mismatch appears in supervisor evaluation. The co-supervision mandate asks local supervisors to support doctoral internationalization through Scopus-indexed publication, international conference participation, and research mobility. Yet these forms of supervisory labor do not appear to carry equivalent weight in the systems through which local supervisors are assessed, promoted, or rewarded. Promotion and performance systems tend to make indexed publication visible through rankings, KPIs, and publication rewards, while the slower labor of doctoral mentorship and local supervisory capacity-building remains weakly recognized (Kuzhabekova & Ruby, 2018; Moldashev & Tleuov, 2022; Kudaibergenova et al., 2022).

Producing a student's indexed publication, preparing a student for a conference presentation, or helping secure a research internship advances the formal goals of the doctoral framework. If these activities remain weakly visible in performance review, the mandate creates obligations without giving local supervisors a

strong institutional reason to prioritize them. The misalignment is not a minor administrative problem. It shapes what supervisors can rationally afford to treat as career-relevant work.

Supervisor evaluation should recognize documented contributions to doctoral students' internationalization outcomes. This does not mean turning students into promotion metrics. It means recognizing the supervisory labor through which students meet requirements that the state and universities have already imposed on them. Indexed student publications, international conference presentations, and research mobility outcomes should count in supervisor performance reviews where there is evidence of substantive supervisory involvement. Without that recognition, the system asks local supervisors to support outcomes that do not meaningfully support their own career progression.

The deeper question concerns the purpose of co-supervision. If internationalization outcomes pass mainly through the foreign supervisory relationship, the system is operating as a delivery mechanism. Students gain access to publication advice, conference networks, and mobility pathways. The doctoral framework records the outputs it seeks. The arrangement has practical value, especially for current doctoral candidates.

Its weakness is accumulation. The gains may remain attached to individual students and their foreign supervisors rather than becoming local supervisory capacity. Altbach and Knight (2007) distinguish internationalization oriented toward capacity building from internationalization oriented toward prestige or output accumulation. By that standard, the long-term test is whether local supervisors become more able, over time, to guide later doctoral cohorts toward international publication, conference participation, and research mobility. Foreign co-supervision would then supplement local capacity. It would not carry the model's core function.

If each cohort requires the same level of foreign supervisory input as the previous one, the arrangement is repeating the dependency it was introduced to address (Shahjahan & Kezar, 2013). A different outcome requires a different object of investment. Local supervisors cannot remain only formal partners in a policy whose internationalizing work is performed elsewhere. They need to become the primary site of the capacity-building effort: through workload reform, funded mobility, recognition in evaluation systems, and explicit opportunities to acquire the publishing and networking knowledge that the doctoral framework now demands.

Limitations

The first limitation concerns the satisfaction measure itself. Satisfaction was measured through an ordinal self-report item, so it records how students judged the supervisory relationship. It does not record what supervisors actually did. The dataset does not show how often feedback was provided, how detailed it was, whether it addressed journal selection or revision strategy, or whether students used that feedback in their publication, conference, or internship work. Satisfaction should therefore be treated as a perception-based indicator, not as a direct measure of supervisory practice.

The cross-sectional design also limits the interpretation of the associations. The analysis cannot establish direction. One possible reading is that stronger supervisory feedback helped students publish, present at conferences, or complete an internship abroad. Another is that students who had already achieved these outcomes rated their supervisors more favorably afterward. A student with a Scopus publication, for example, may look back on the supervisory relationship differently from a student still struggling to submit a manuscript. The data cannot separate these possibilities.

There is also unmeasured context around the student. Institutional affiliation, prior research preparation, access to English-language support, and familiarity with international publishing may shape both satisfaction ratings and internationalization outcomes. These conditions may not be evenly distributed across doctoral programs or supervisory arrangements. Some of the observed differences between local and foreign supervisor satisfaction may therefore reflect wider institutional and linguistic resources rather than supervision alone.

For these reasons, the findings should be read as evidence of patterned association. They show that satisfaction with local and foreign supervisors relates differently to publication, conference participation, and internship outcomes. They do not show why those differences occur. Nor do they establish whether the observed patterns result from supervisory influence, student characteristics, institutional resources, or some combination of these conditions.

Conclusion

The findings show a clear asymmetry in Kazakhstan's co-supervision arrangement. Satisfaction with foreign supervisor feedback is associated with all three doctoral internationalization outcomes examined in this study: publication output, conference participation, and internship outcomes. Satisfaction with local supervisor feedback is associated only with internship outcomes, and even there, the relationship is weaker.

This pattern should not be read as evidence that foreign supervisors are simply better supervisors. The more plausible explanation is structural. Local supervisors are expected to contribute to the same internationalization outcomes as their foreign counterparts, but they often work under less favorable conditions for doing so. Heavy teaching and administrative loads reduce the time available for manuscript development and sustained doctoral feedback (OECD, 2017; Kuzhabekova & Ruby, 2018; Moldashev & Tleuov, 2022). More limited access to international mobility and research collaboration restricts the networks through which publication, conference participation, and external research opportunities are usually built (Kuzhabekova, Ispambetova, et al., 2022; Kuzhabekova, Bilyalov, et al., 2022). Kazakhstan's doctoral framework designates local and foreign supervisors as equal partners. The evidence from this study suggests that the partnership does not operate on equal terms.

The policy problem is therefore not only supervisory coordination. It is capacity. If doctoral internationalization runs mainly through the foreign supervisor, the model can still produce visible gains for current students. It can help them publish, attend conferences, and complete research stays abroad. But those gains remain fragile if they depend on repeated access to foreign supervisory expertise rather than on the growth of local supervisory capacity.

For co-supervision to become more than a delivery mechanism, local supervisors need the conditions under which internationalizing supervision can be performed. That means reduced teaching loads, funded mobility, access to international research networks, and evaluation systems that recognize doctoral mentorship as career-relevant academic labor. Without those changes, the co-supervision requirement will continue to produce the formal appearance of partnership while leaving the substance of that partnership unevenly distributed.

Funding Information

This study was funded by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan under Grant No. AP23488831, "Reforms in Kazakhstani Higher Education through Internationalization and Transformational Leadership."

Conflict of Interest Statement

We declare no potential conflicts of interest with respect to the research, authorship, or publication of this article.

Author Contributions

Askat Tleuov, Mukhametzhan Seitzhapparuly, and Kairat Moldashev contributed equally to the conceptualization, methodology, data analysis, interpretation of the findings, preparation of the original manuscript, review, editing, and finalization of the study. All authors participated equally in the research and approved the final version of the manuscript.

Declaration of Use of Generative AI and AI-Assisted Technologies

The authors declare that no generative AI or AI-assisted technologies were used in the preparation of this manuscript.

References

- Altbach, P. G., & Knight, J. (2007). The internationalization of higher education: Motivations and realities. *Journal of Studies in International Education*, 11(3–4), 290–305. <https://doi.org/10.1177/1028315307303542>
- Dericks, G. H., Thompson, E. R., Roberts, M., & Phua, F. T. T. (2019). Determinants of PhD student

- satisfaction: The roles of supervisor, department, and peer qualities. *Assessment & Evaluation in Higher Education*, 44(7), 1053–1068. <https://doi.org/10.1080/02602938.2019.1570484>
- Halse, C., & Malfroy, J. (2010). Retheorizing doctoral supervision as professional work. *Studies in Higher Education*, 35(1), 79–92. <https://doi.org/10.1080/03075070902906798>
- Hauss, K. (2021). What are the social and scientific benefits of participating at academic conferences? Insights from a survey among doctoral students and postdocs in Germany. *Research Evaluation*, 30(1), 1–12. <https://doi.org/10.1093/reseval/rvaa018>
- Kalman, O., Horvath, L., Kardos, D., Kozma, B., Feyisa, M. B., & Ronay, Z. (2022). Review of benefits and challenges of co-supervision in doctoral education. *European Journal of Education*, 57(3), 497–511. <https://doi.org/10.1111/ejed.12518>
- Kehm, B. M., & Teichler, U. (2007). Research on internationalization in higher education. *Journal of Studies in International Education*, 11(3–4), 260–273. <https://doi.org/10.1177/1028315307303534>
- Khuram, W., Wang, Y., Ali, M., Khalid, A., & Han, H. (2023). Impact of supportive supervisor on doctoral students' research productivity: The mediating roles of academic engagement and academic psychological capital. *SAGE Open*, 13(3), 1–15. <https://doi.org/10.1177/21582440231185554>
- Kudaibergenova, R., Uzakbay, S., Makanova, A., Ramadinkyzy, K., Kistaubayev, E., Dussekeev, R., & Smagulov, K. (2022). Managing publication change at Al-Farabi Kazakh National University: A case study. *Scientometrics*, 127(1), 453–479. <https://doi.org/10.1007/s11192-021-04139-y>
- Kuzhabekova, A. (2025). Ph.D. publication requirement and its effects on research productivity trends in Kazakhstan. *Higher Education Quarterly*, 79(1), e12590. <https://doi.org/10.1111/hequ.12590>
- Kuzhabekova, A., Bilyalov, D., & Mussabayeva, M. (2022). Faculty views on internationalization in Kazakhstan. *Yükseköğretim Dergisi*, 12(Supplement), 35–44. <https://doi.org/10.2399/yod.21.202204>
- Kuzhabekova, A., Ispambetova, B., Baigazina, A., & Sparks, J. (2022). A critical perspective on short-term international mobility of faculty: An experience from Kazakhstan. *Journal of Studies in International Education*, 26(4), 454–471. <https://doi.org/10.1177/10283153211016270>
- Kuzhabekova, A., & Ruby, A. (2018). Raising research productivity in a post-Soviet higher education system: A case from Central Asia. *European Education*, 50(3), 266–282. <https://doi.org/10.1080/10564934.2018.1444942>
- Kuzhabekova, A., Sparks, J., & Temerbayeva, A. (2019). Returning from study abroad and transitioning as a scholar: Stories of foreign PhD holders from Kazakhstan. *Research in Comparative and International Education*, 14(3), 412–430. <https://doi.org/10.1177/1745499919868644>
- Lee, A. (2008). How are doctoral students supervised? Concepts of doctoral research supervision. *Studies in Higher Education*, 33(3), 267–281. <https://doi.org/10.1080/03075070802049202>
- Lee, A., & Green, B. (2009). Supervision as metaphor. *Studies in Higher Education*, 34(6), 615–630. <https://doi.org/10.1080/03075070802597168>
- Mainhard, T., van der Rijst, R., & Wubbels, T. (2009). A model for the supervisor–doctoral student relationship. *Higher Education*, 58(3), 359–373. <https://doi.org/10.1007/s10734-009-9199-8>
- Manathunga, C. (2014). *Intercultural postgraduate supervision: Reimagining time, place and knowledge*. Routledge. <https://doi.org/10.4324/9781315818854>
- McAlpine, L., & McKinnon, M. (2013). Supervision—the most variable of variables: Student perspectives. *Studies in Continuing Education*, 35(3), 265–280. <https://doi.org/10.1080/0158037X.2012.746227>
- Moldashev, K., & Tleuov, A. (2022). Response of local academia to the internationalization of research policies in a non-Anglophone country. *Education Policy Analysis Archives*, 30(56). <https://doi.org/10.14507/epaa.30.6788>
- Nerad, M., & Evans, B. (Eds.). (2014). *Globalization and its impacts on the quality of PhD education: Forces and forms in doctoral education worldwide*. Springer.
- OECD. (2017). *Higher education in Kazakhstan 2017*. OECD Publishing. <https://doi.org/10.1787/9789264268531-en>
- Orellana, M. L., Darder, A., Pérez, A., & Salinas, J. (2016). Improving doctoral success by matching PhD students with supervisors. *International Journal of Doctoral Studies*, 11, 87–103. <https://doi.org/10.28945/3404>
- Sagintayeva, A., & Kurakbayev, K. (2015). Understanding the transition of public universities to institutional

- autonomy in Kazakhstan. *European Journal of Higher Education*, 5(2), 197–210. <https://doi.org/10.1080/21568235.2014.967794>
- Shahjahan, R. A., & Kezar, A. J. (2013). Beyond the “national container”: Addressing methodological nationalism in higher education research. *Educational Researcher*, 42(1), 20–29. <https://doi.org/10.3102/0013189X12463050>
- Shen, W., & Jiang, J. (2023). Institutional prestige, academic supervision and research productivity of international PhD students: Evidence from Chinese returnees. *Journal of Sociology*, 59(2), 552–579. <https://doi.org/10.1177/14407833211055225>
- Sverdlik, A., Hall, N. C., McAlpine, L., & Hubbard, K. (2018). The PhD experience: A review of the factors influencing doctoral students' completion, achievement, and well-being. *International Journal of Doctoral Studies*, 13, 361–388. <https://doi.org/10.28945/4113>
- Tokatligil, Y., Saliyeva, A., Mambetalina, A., Nurmanbetova, J., & Karmelyuk, A. (2025). Co-supervision in Kazakhstan PhD programs: The Kazakhstani international co-supervisor model. *International Journal of Doctoral Studies*, 20, Article 15. <https://doi.org/10.28945/5626>
- Vattøy, K. D. (2026). The role of international research stays in the professional development of doctoral students in teacher education: Impacts on qualification, socialisation, and subjectification. *Teachers and Teaching*, 32(2), 145–162. <https://doi.org/10.1080/13540602.2026.2642687>
- Wilkins, S., Hazzam, J., & Lean, J. (2021). Doctoral publishing as professional development for an academic career in higher education. *The International Journal of Management Education*, 19(1), 100459. <https://doi.org/10.1016/j.ijme.2021.100459>
- Wu, Y., Chen, Z., & Zhang, D. (2025). Supervisors' academic supervising behaviors and graduate students' academic thriving: The mediating roles of admiration and task engagement. *Behavioral Sciences*, 15(6), Article 754. <https://doi.org/10.3390/bs15060754>
- Yelibay, M., Karabassova, L., Mukhatayev, Z., & Yermukhambetova, A. (2022). The perception and experience of young researchers in doctoral programmes in the context of recent reforms in Kazakhstan. *European Journal of Education*, 57(3), 484–496. <https://doi.org/10.1111/ejed.12511>

Information about authors

Askat Tleuov – KIMEP University, College of Human Sciences and Education, Republic of Kazakhstan; e-mail: a.tleuov@kimep.kz; ORCID: 0000-0001-6642-3047 (**corresponding author**).

Mukhametzhana Seitzhapparuly – Narxoz University, Republic of Kazakhstan; e-mail: mukhametzhana.seitzhapparuly@narxoz.kz; ORCID iD: 0009-0007-5820-9308.

Kairat Moldashev – SDU University, Republic of Kazakhstan; e-mail: kairat.moldashev@sdu.edu.kz; ORCID: 0000-0001-6558-6004.

Received: 29.06.2026

Revised: 17.08.2026

Accepted: 25.09.2026

Published: 30.09.2026

Type of Article: Research Article

¹Magzhan Amangeldi, ^{1,2}Oleksandr Kuchanskyi*

¹ Astana IT University, Astana, Republic of Kazakhstan

² National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

FROM CITATION COUNTS TO COMPUTATIONAL SCIENCE: METHODS, DATA, AND APPLICATIONS IN SCIENTOMETRICS

Abstract. In the last 60 years, the science of scientometrics, the quantitative analysis of science as a communication system, has developed from an esoteric interest in citation indexing to a key tool for research policy and evaluation. The aim of this review is to trace and summarise that trajectory under the single argument that scientometric indicators are only as meaningful as the data infrastructures on which they are built, and the evaluative contexts in which they are applied. We present the literature in five methodological layers, linked by a common thread on research evaluation and policy, and spanning bibliographic data sources, citation indicators and normalisation, science mapping and visualisation, altmetrics and the widening of impact, as well as the recent computational turn based on machine learning and natural language processing. Along with this narrative review, we present a companion bibliometric analysis, mapping the field by topic within the OpenAlex database, which reveals an analytic corpus of 37,166 works, whose growth we illustrate, whose core journals we identify, and whose thematic structure we discuss as a characterization of the field from a methodological perspective. The review finds that the best lasting gains in the field have been in the quality of data coverage, data normalisation and in the responsible interpretation of indicators, and the current transition from counting to understanding indicators should be subject to the same standards.

Keywords: scientometrics, bibliometrics, citation analysis, research evaluation, data infrastructures, openalex, science mapping, altmetrics, machine learning, natural language processing, citation normalization, open science, knowledge graph, scholarly communication, computational scientometrics.

Introduction

The quantification of research evaluation has become a phenomenon that has taken root over the last 20 years. The number of publications, citations and their variants are becoming more and more important in the decision-making process for hiring and promotion, in the distribution of grants to research institutions, in the ranking of universities and in the comparison of national research systems (Faghri & Bergman, 2024; Hicks et al., 2015).

This change has led to the emergence of scientometrics as an information science discipline with tangible impact on the funding, organization, and reward of science. Scientometrics should be considered as an interdisciplinary subject which focuses on the quantitative aspects of science as a communication system, and it is closely but not exclusively related to the analysis of citations (Leydesdorff et al., 2016). It is related to, but not the same as the neighbouring areas of bibliometrics and informetrics (Ellegaard, 2026; Hood & Wilson, 2001).

In practice, though, lines are blurred, and in recent years the discourse has grown more about a common science of science that incorporates macro-level aspects of scholarly production, along with models from network science, economics, and machine learning (Fortunato et al., 2018). The scholarly trajectory of the discipline can be understood as a series of extensions. It starts with Garfield's proposal for a citation index for the sciences, making the citation a tractable machine-countable measure of scholarly influence (Garfield, 1955) and its most significant individual-level instrument is Hirsch's h-index (Hirsch, 2005).

The field has then in turn evolved from raw counts to normalised indicators (Ioannidis et al., 2016; Lu et al., 2026; Waltman & van Eck, 2019); indicators to maps of the structure of science (Klavans & Boyack, 2017; Van Eck & Waltman, 2017); citations to altmetrics (González et al., 2025; Sugimoto et al., 2017); and finally, counting citations to understanding citations (Jurgens et al., 2018; Saeidnia et al., 2025). For this review that path is the organising spine. Under these expansions, there's a continuous feeling of tension. Each indicator is calculated in some database, and databases vary systematically in coverage (Culbert et al., 2025; Mongeon

& Paul-Hus, 2016); each indicator is understood in some evaluative regime, and the regimes influence the behaviour they claim to measure (Baccini et al., 2019). The Leiden Manifesto and the San Francisco Declaration on Research Assessment were both born out of the misuse and overuse of metrics (Gagliardi et al., 2023; Hicks et al., 2015) and there is significant literature warning that citations cannot be taken as a simple indicator and must be understood within their context (Aksnes et al., 2019; Ioannidis et al., 2016).

This paper provides a structured synthesis that is organized around four questions: (i) What are the data infrastructures behind the indicators of scientometrics, and how do the indicators of different data infrastructures differ in terms of their coverage and bias; (ii) How have the indicators of scientometrics and their normalisation changed; (iii) How do science-mapping methods reconstruct the structure of research; and (iv) How is the computational turn reshaping what scientometrics can measure. We are not tied to a national system and have a general scope of work. While reviews focus on a certain aspect of the field such as indicators (Waltman & van Eck, 2019), software tools (Moral-Muñoz et al., 2020), or AI methods (Safder et al., 2023), we combine across the 5 layers and we link the five layers together through the evaluation thread, and we supplement the review with a companion analysis based on an open reproducible data source (Donthu et al., 2021).

Materials and Methods

Review design. This paper follows a hybrid design: the structured narrative review of the methods, transparent and explicit in a selection protocol, and a companion bibliometric analysis of a larger harvested corpus. The narrative layer provides interpretation, the bibliometric layer provides quantitative evidence of growth, structure and trends (Donthu et al., 2021). To make the points above clearer, Table 1 compares the approaches to reviewing found here. The narrative component was question-led rather than designed as a systematic effectiveness review or meta-analysis. Sources were retained when they directly addressed one of the five methodological layers, provided methodological or empirical evidence relevant to research evaluation, or represented a landmark statement, dataset, or infrastructure. Recent review articles were used to identify the principal methods within each layer, after which backward citation tracing was used to recover foundational studies. The search and selection process continued through June 2026, prioritised peer-reviewed international literature, and retained policy statements when they directly shaped responsible-metrics practice. These procedures provide transparent boundaries for the narrative synthesis while the companion bibliometric analysis supplies reproducible descriptive evidence at corpus level.

Table 1

Comparison of review types relevant to scientometric synthesis

Dimension	Narrative	Systematic (PRISMA)	Bibliometric	Meta-analysis
Primary goal	Interpretive synthesis	Answer a focused question	Map structure and trends	Pool effect sizes
Reproducibility	Low	High	High	High
Breadth	High	Narrow, deep	Very high	Narrow
Content meaning	Strong	Strong	Weak	Moderate
This paper	Core method	Protocol adopted	Companion analysis	Not used

Corpus construction. The companion analysis is based on OpenAlex, an open and fully reproducible bibliographic database, selected because its open API allows the corpus to be re-harvested, and because it is covered by a comparable number of publications to established databases (Visser et al., 2021). We segment the field into topics rather than free-text queries: we retrieve works with a major topic on OpenAlex of ‘Scientometrics and Bibliometrics Research’ (OpenAlex topic T10102) or ‘Web Visibility and Informetrics’ (OpenAlex topic T13976) and have been published in English in the time interval 2000–2025. Topic

delineation is preferred over keyword queries because it means that it reduces contamination by the search terms used to define the query and the classification itself acts as a relevance filter. Every query and the snapshot of the corpus are stored along with the date it was retrieved.

For reproducibility, the two topic identifiers were fixed in the retrieval script rather than resolved anew at each run. Works were retrieved through the OpenAlex /works endpoint on 4 June 2026, restricted to English-language articles and reviews published between 2000 and 2025. The outer-boundary query, in which either topic appeared in topics.id, returned 76,369 works. Requiring either T10102 or T13976 in primary_topic.id retained 38,319 works before venue cleaning. A venue-level audit then identified exactly 1,153 misclassified records from five OpenAlex source IDs: C&EN Global Enterprise (S4210177211, n = 56), Chemical & Engineering News (S163506364, n = 11; S4210207877, n = 814), Physics Today (S180907581, n = 155), and Physics World (S9810779, n = 117). These records were removed once, before every analysis. The resulting final cleaned analytic corpus contained 37,166 works, and this same corpus was used for annual output, Bradford analysis, open-access analysis, topic-stability assessment, and term co-occurrence. The complete API response metadata and the retrieved work-level fields were retained locally so that the filtering and cleaning sequence could be reproduced.

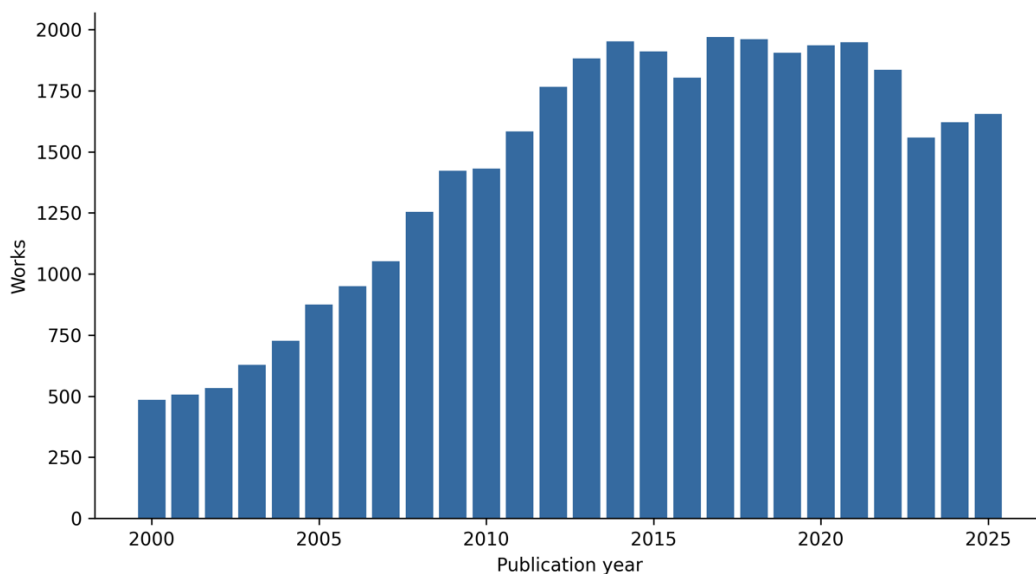
Corpus adequacy. Convergent criteria: saturation of the annual output, Bradford core-zone concentration, structural stability of topic composition, and a venue-level audit of anomalously dense clusters were used to assess the adequacy of the corpus size rather than merely asserting its adequacy. Because the corpus is defined by OpenAlex topic classification rather than keyword matching, its validity depends on the quality of that classification. The venue audit detected the five misclassified source IDs listed above, and the same exclusion rule was then applied globally. No independent hand-coded reference corpus was available; therefore, the study does not claim a formal recall estimate or exhaustive coverage. Instead, validation addresses observable classification noise, the stability of the principal topic structure, and the consistency of the final corpus across all reported analyses. This limitation is stated explicitly so that corpus adequacy is not confused with perfect topic-classification accuracy.

Open-access status was obtained from the OpenAlex field open_access.is_oa. Records coded true were treated as open access and records coded false as closed access, aggregating the individual OA routes for the open-versus-closed comparison. There were no missing is_oa values among works published from 2000 through 2024. Shares were calculated within consecutive five-year periods using the final cleaned corpus as the denominator for each period. The incomplete publication year 2025 was retained in the overall corpus and annual-output analysis but excluded from the five-year OA comparison. Under this procedure, the OA share increased from 20.7% in 2000–2004 to 68.0% in 2020–2024.

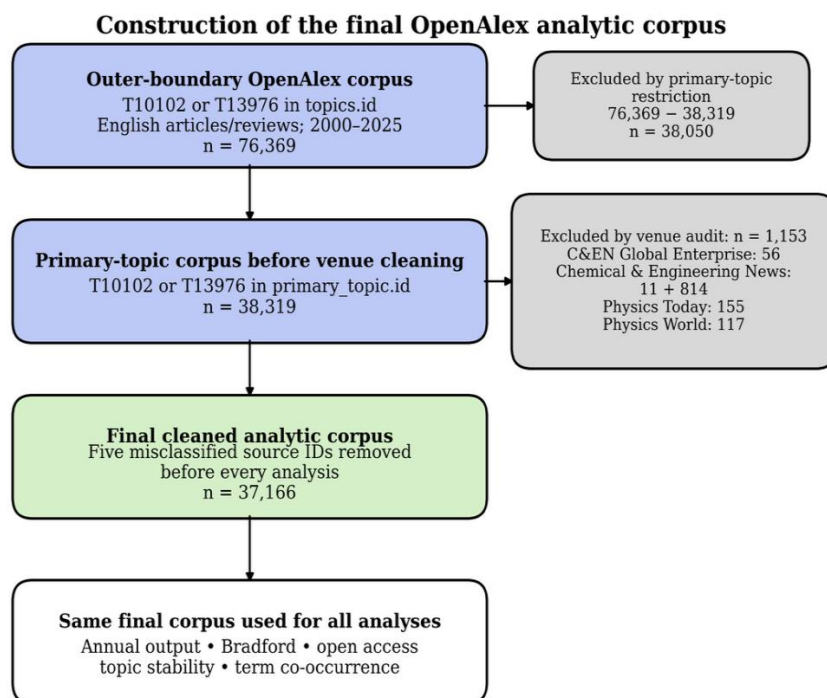
Results of corpus construction. The two-topic outer boundary contained 76,369 records, of which 38,319 assigned T10102 or T13976 as the primary topic. The venue audit excluded 1,153 misclassified trade and news records, leaving a final cleaned analytic corpus of 37,166 works. The distinction between the outer boundary and the analytic corpus remains important: many outer-boundary works apply a bibliometric method to a substantive domain, whereas the primary-topic corpus focuses on work classified as a methodological contribution to scientometrics or informetrics. Annual output increased from 486 works in 2000 to a maximum of 1,970 in 2017 and remained close to 1,800–1,950 during much of 2018–2022; lower counts in 2023–2025 may partly reflect indexing lag. The Bradford analysis included 27,166 works assigned to journal sources. Its first zone comprised 37 sources and 9,095 works, or 33.5% of the Bradford denominator. The most prominent sources included Scientometrics, Journal of Informetrics, Journal of the Association for Information Science and Technology, and Research Evaluation. Figure 2 summarises the identification, primary-topic restriction, venue cleaning, and final inclusion stages.

Figure 1

Annual output of the final cleaned analytic corpus ($n = 37,166$), 2000–2025. Output peaks at 1,970 works in 2017; the lower recent counts may partly reflect incomplete indexing.

**Figure 2**

Flow diagram for construction of the single-source, topic-delineated corpus. The outer-boundary set contained 76,369 records; the primary-topic restriction retained 38,319; venue cleaning removed 1,153; and the final analytic corpus contained 37,166 works.



Retrieval date: 4 June 2026. Counts refer to records after each sequential restriction.

Results and Discussion

Data sources and infrastructure. Each of the scientometric indicators is calculated with respect to some database, and no database is a neutral window on science. In the end, coverage decisions regarding journals, languages, document types, and years shape the indicators they are used to create, and the choice of a database

is a methodological decision, not a technical preliminary. For 40 years, big bibliometrics was based on just one source: the Science Citation Index (and later Web of Science, WoS); with the introduction of Scopus in 2004, the monopoly was broken and the comparison of databases became a research question.

The two are the primary reference databases for evaluating research, but they are not synonymous. Both databases demonstrate that the natural sciences, engineering, and biomedical research are overrepresented compared to the social sciences and the arts and humanities in their coverage of journals; both also favour publications in English over those in other languages, and, while the two databases cover different journals and the coverage of each journal varies significantly, the choice between the two can affect the outcome of a journal analysis (Mongeon & Paul-Hus, 2016).

A third, more widely used, source is Google Scholar (GS) which covers a much wider, and less curated, selection of content, such as books, conference papers, and work not written in English. The choice among these data sources and metrics affects the results of cross-disciplinary comparisons (Abramo et al., 2026). The lesson that runs through this literature is that bibliometric results depend on the source itself – an indicator is only as comparable as the coverage that it is based on.

Since then, the volume of data has expanded significantly. Launched in 2018, Dimensions goes beyond the publication–citation core and extends its reach to include grants, patents and clinical trials, as first-class objects that are connected within one graph and are linked with institutions and research classifications by machine learning (Hook et al., 2018). Crossref, the DOI registration agency, has become a significant open source of scholarly metadata in its own right, with over 100 million records not only of journal articles, but of datasets, preprints, peer reviews, grants, and more, and linking increasingly to citations, abstracts, and funding.

The most systematic assessment of this expanded field is a comparison of the coverage, completeness, and accuracy of citation links across the five sources Scopus, WoS, Dimensions, Crossref and Microsoft Academic on a large scale, which suggests that the practical implication is to integrate comprehensive coverage with flexible filtering, not to rely on any single source as a primary authority (Visser et al., 2021). Independent analysis of the literature covered in Google Scholar agrees: Google Scholar is able to find a very high proportion of the open-access literature, but it is not as consistent or convenient as it would be to measure open-access literature systematically (Martín-Martín et al., 2018; Weimer et al., 2023).

A separate strand of this literature is to do with control, not coverage. The analyses using WoS and Scopus are difficult to reproduce, as these are proprietary, selectively indexed, and biased against the humanities and social sciences, non-English languages, and research from non-Western countries. The open-infrastructure response, in the form of Crossref's community-owned metadata (Hendricks et al., 2020) and the OpenCitations initiative, and the latest OpenAlex, are all designed to offer a nonproprietary, fully queryable substrate for scientometric research. It's not only political – because having an open, versioned source means that someone can re-execute it in principle, which is the requirement that the evaluation literature imposes on the metrics that it's been critiquing.

This review has a high regard for that standard in its design. The companion analysis is developed specifically using the open API of OpenAlex, which also allows the entire corpus to be re-harvested, and for the figures to be re-generated, while also providing citations and coverage that compare favourably with the established sources (Haunschild & Bornmann, 2024; Visser et al., 2021). But the compromise is real: open sources can be noisier, as the experience we had with several misclassified high-volume trade and news publications revealed (Zhang et al., 2024); they also need to be clearly cleaned and differentiated. That cost, however, gives transparency: All selection decisions are transparent and reversible, a proprietary black box cannot provide that.

The main take-home message for any scientometric study is that the data source must be selected carefully and documented completely; the source is not merely a matter of convenience, it is a key determinant of the result. Citation indicators, normalization. The indicator is not the data source; it sets the meaning of the count that is provided by the data source. One of the fundamental problems of scientometrics is that raw citation counts are not comparable: a paper in chemistry and a paper in mathematics with the same number of citations may have very different impact scores.

Much of the literature on citation impact indicators is the history of the efforts to make the counts comparable and of warnings about how the comparisons can mislead. The most basic single indicator is the h-index (Hirsch, 2005), which is the maximum value such that a researcher has *h* papers that have been cited *h*

or more times. It is an intuitive number that condenses productivity and impact in one, and this sparked its quick adoption across hiring, promotion, and funding decisions. This also its strength: the h-index is not affected by outliers, meaning that its value grows with the length of the career, and like any raw count — it is dependent on the field, which means that there are many variants and field normalised versions (Lu et al., 2026).

Indicators surveys emphasize that there is no single indicator, and point out that there needs to be a mix of quantitative indicators and qualitative judgment, as well as choosing indicators suitable for the evaluative question under consideration (Faghri & Bergman, 2024). A wide range of reviews of other citation impact indicators groups the field around a handful of common challenges: what publications and citations to count, how to distribute credit for co-authorship, how to build journal-level indicators, and, most importantly, how to adjust for field differences (Waltman & van Eck, 2019).

The technical core of modern bibliometrics is normalization. The typical approach involves comparing a paper's citations to that of the expected number of citations for the field and publication year, but there is debate over how to make that expectation. These scores could be normalised by journal subject categories or, more defensibly, by classification systems that categorise each publication into a main research area prior to calculating percentile-based normalised scores, with percentile approaches being robust to the skew in citation data, and avoiding the distortions of mean-based ratios (Lu et al., 2026).

However, the issue of normalisation is not resolved. every normalisation embeds additional assumptions: what constitutes a field boundary, what constitutes a document type, how much the databases cover, and so on, and further complications can be introduced in cases of aggregation of indicators to evaluate scientists, institutions, journals, or countries, and in cases where credit has to be given in the context of multiauthored work (Aksnes et al., 2019; Ioannidis et al., 2016). The guidance is not prescriptive; different indicators provide a different perspective and the assumptions behind each indicator need to be explained and not hidden within the composite score.

The journal impact factor is the most often-used and criticized factor, and it is the use of the impact factor as a proxy for the quality of individual articles that is one of the main criticisms of the research-assessment reform movement (Gagliardi et al., 2023; Hicks et al., 2015). At the level of countries and institutions, larger dimensions are given a higher status and it is recommended to use the size-dependent indicators in combination with the size-independent ones; for example, an empirical study of the Nature Index suggests that the size-dependent version systematically favours large countries and institutions, providing less information about the performance of small producers, while simpler versions reproduce a lot of the information provided by the Nature Index (Bornmann & Haunschild, 2017).

The general rule is similar to that found in the data-source literature: an aggregate indicator is an interpretive instrument that, to serve its purpose, must be constructed to reflect the question being asked, with the limitations of the instrument reported along with its values. Science mapping and visualization. The methods discussed above count papers, authors and journals. But science mapping raises a new question: What is the structure of a research field, and how does it change? Mapping techniques, which depict the literature as a network (of co-cited references, of bibliographically coupled documents or of co-occurring terms) and divide that network into clusters, depict the intellectual organisation of a field, which no other indicator can (Klavans & Boyack, 2017; Van Eck & Waltman, 2017). There are three relational signals that are well represented in the literature, and the selection of one or another is significant.

Co-citation analysis is performed if two earlier documents are both cited in a later one; often documents that are co-cited are interpreted as having a shared intellectual origin. Bibliographic coupling connects two documents that share references, instead of similarity of reception, it represents similarity of intellectual foundations. Direct citation connects a citing document to the cited document. Direct citation has been shown to yield the most accurate topic-level taxonomy of science in the most systematic comparison to date, where the direct citation-based taxonomy is also closer to focusing references than the other two approaches (bibliographic coupling, co-citation), and where use of discipline-level maps constructed from journal schemes is cautioned as being significantly less accurate than document-level maps (Klavans & Boyack, 2017).

The fourth and more recently popular strategy is to skip the citations entirely: keyword or term co-occurrence, in which nodes are terms, and edges are formed between the nodes that co-occur in documents, capturing the conceptual structure of a field instead of the citation structure (Radhakrishnan et al., 2017). The

advantages of term co-occurrence are that only titles, abstracts or author keywords are needed to create a concept, but it also suffers from the ambiguities of natural language: synonymy, polysemy and the presence of generic terms and is therefore very sensitive to term cleaning. A surge of mapping studies is not the result of new methods, but more of tools that are accessible.

Comparative reviews include a stable division of labour: the visualization of large networks using distance-based matrices (VOSviewer); the identification of emerging trends in terms of term bursts (CiteSpace); and the broadest set of performance and mapping analyses (R package bibliometrix); and the identification of emerging trends in terms of terms and bursts (SciMAT). A recent meta-bibliometric evaluation of over ten thousand studies, carried out using VOSviewer, revealed that the tool is used almost everywhere and a growing methodological tradition focused on citation- and keyword co-occurrence analysis, as well as some frequent methodological weaknesses, such as relying on one database, keyword metadata that varies from study to study and the sensitivity to threshold values (Genc & Kocak, 2024; Omer et al., 2025). The main weakness lies not accidentally in the science map, it is a main weakness; it is the main motivation for the methodological transparency used below.

We created a term co-occurrence map on the analytic corpus to illustrate the method and describe the conceptual network of the field. Instead of author keywords, the terms were mined from the title and abstract of the article, since the keywords in open data sources are not always present and are often inconsistent; the same text-mining approach as used natively by VOSviewer. The 400 common terms used by the four most common sources were not excluded, while an explicit, reported stopword list excluded generic academic vocabulary or source-specific boilerplate. Text was converted to lowercase, non-alphabetic characters were removed, and whitespace was normalised. Binary document-term occurrences were generated with scikit-learn CountVectorizer using one- to three-word n-grams, a minimum document frequency of 40, the standard 318-entry English stopword list, and 170 additional domain-specific entries; after overlap, the combined list contained 469 unique stopwords. Of 6,293 candidate terms meeting the frequency threshold, the 400 most frequent were retained.

Documents containing the same pair of retained terms were linked, and edge strength was defined as the number of documents in which both terms occurred. Links with strength below 5 were excluded, leaving 79,742 links among 400 retained terms derived from all 37,166 documents in the final cleaned corpus. The resulting map and network files were imported into VOSviewer 1.6.20. Association-strength normalisation was used; layout attraction was 2 and repulsion was 0; clustering resolution was 1.00, the minimum cluster size was 1, and random seed 42 was used. Node size represents document occurrence, link strength represents document-level co-occurrence, and colour represents the VOSviewer cluster assignment. We explicitly report the stopword policy because, as the software literature cautions (Omer et al., 2025), term cleaning is a major source of investigator discretion. The initial map exposed a dense peripheral cluster of metadata fragments associated with the five misclassified source IDs. Removing all 1,153 records before every analysis eliminated that contamination and illustrates the direct propagation of data-source quality into mapping results.

The map in Figure 3 divided into four main clusters. The methodological nucleus of citation analysis consists of three major groups: citation, bibliometric, impact factor, the h-index, productivity, and the big databases, along with the vocabulary of mapping and ranking. A second large cluster is the open access, scholarly communication and research policy: open access to publishing to funding to institutional to policy to visibility and to research community. The third compact dimension that the map recovers as a region rather than a subset of citation impact is that of social, online, attention, media and “alternative” – this is a grouping that the map allows for visually, and that fits with the empirical observation that altmetrics provide a different signal from citations (González et al., 2025; Sugimoto et al., 2017). A fourth, peripheral cluster relates to peer-review and publishing process: peer review, editors, reviewers, conferences, proceedings, acceptance, and predatory publishing.

The spatial relations between the clusters are as informative as the clusters themselves: between the citation core and the open-access cluster lies the altmetrics cluster, which are mediated by bridge terms like “social”, “attention” – the online attention is near the citation impact and the openness of dissemination. Two cautions apply. First, the clusters represent co-occurrence of language, rather than citation-based intellectual lineage: a co-citation map or a direct-citation map would emphasize foundational work and schools of thought, and be more useful for a history-of-ideas reading (Klavans & Boyack, 2017). Second, the map cannot be

considered as an absolute result or as independent of the thresholds and of the stop word selection given above, but as a transparent, reproducible characterisation of the conceptual structure of the field.

Figure 3

Term co-occurrence map of the final cleaned analytic corpus (n = 37,166 documents; 400 retained terms; minimum document frequency = 40; minimum link strength = 5; 79,742 retained links). VOSviewer 1.6.20 settings: association-strength normalisation, attraction = 2, repulsion = 0, clustering resolution = 1.00, minimum cluster size = 1, and random seed = 42. Four clusters represent citation methods and indicators, open access and research policy, altmetrics and online attention, and peer review and publishing.

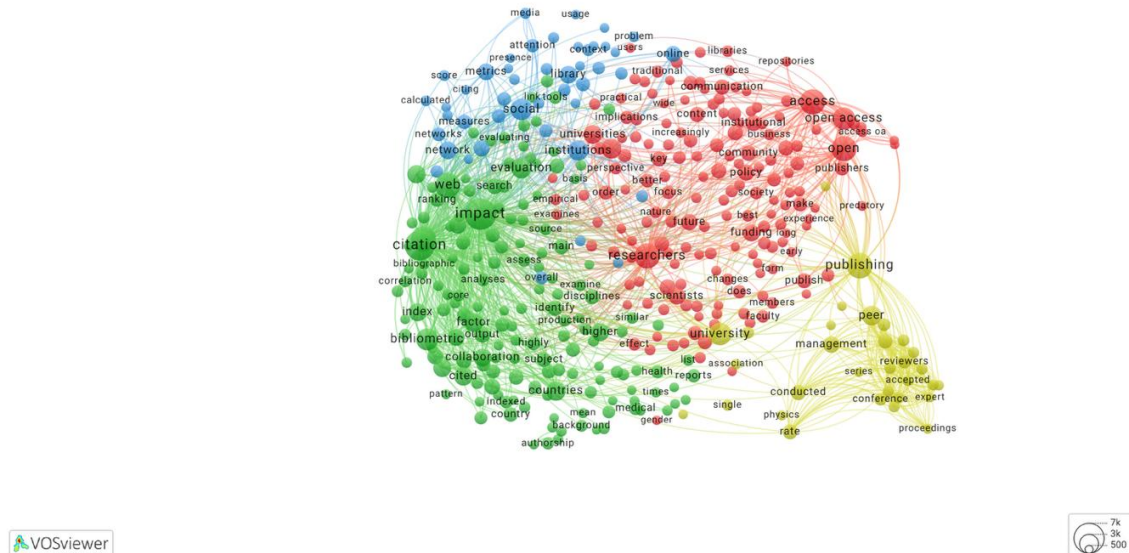
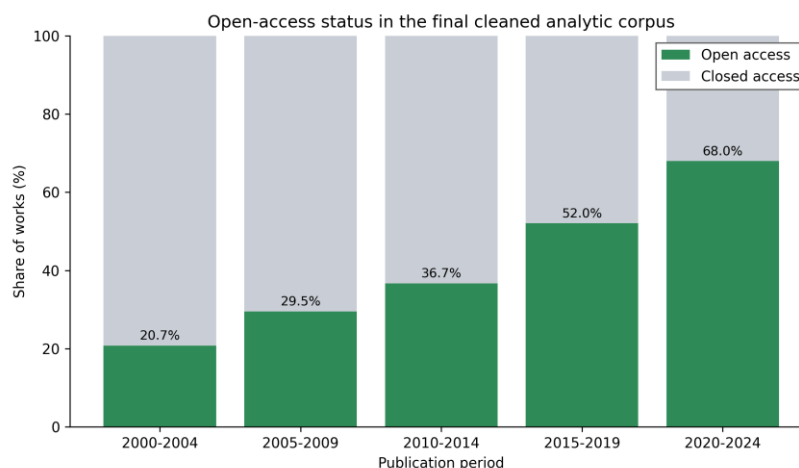


Figure 4

Open-access versus closed-access works in the final cleaned analytic corpus, by five-year period. Open access is defined from OpenAlex `open_access.is_oa`; no values were missing in 2000–2024. The open-access share rises from 20.7% in 2000–2004 to 68.0% in 2020–2024.



Expansion of impact and altmetrics. All of the indicators mentioned above quantify impact in the scholarly literature. Beginning in the early 2010s, a parallel current has emerged that sought to quantify impact outside of it, using the term 'altmetrics' for social media, reference managers, news outlets, policy documents and online platforms. This was for two reasons: firstly to collect influences beyond citations and secondly to help to record impact more quickly than the long timeframe of citations accumulating. This domain was recovered as a distinct cluster; it is a clearly different area from citation analysis within the area of altmetrics. The fundamental review is based on two interrelated phenomena: the use of social media platforms as a

scholarly tool and the indicators based on their interactions.

The distinction is important because it addresses two issues: an issue of scholarly practice, and an issue of measurement validity, and it is important not to mix the two in the literature. Whether or not there is a correlation between the number of mentions on social media and the number of citations has been the central empirical question. The conclusion from a meta-analysis is that they don't really, the correlation between citations and microblogging is negligible, with blog mentions it is small, and with bookmark counts from online reference managers it is medium to large (González et al., 2025). The interpretation of the data has been well-established after some major shifts and the "attention" is now widely agreed to measure attention, not the impact of the scholarly work; impact is linked, but not identical, and the link's strength will vary greatly by platform. Unlike citation data, altmetric data can have measurement difficulties.

The counts are skewed and zero-inflated: most papers are cited on very few platforms, so that mean-based indicators would be misleading, while they require methods with sparse data, such as indicator families that could differentiate between the quality levels which are assessed by citing them on platforms that have been validated by peer assessment, and on those that have not. In addition to sparsity, the sources of altmetrics are unstable and unevenly covered: platforms come and go, coverage is uneven across disciplines, and the sources of counts are often not clearly stated. For example, an analysis of Publons metrics revealed that there are important coverage biases by discipline and publisher, and weak, non-significant relationships between bibliometric and altmetric metrics, meaning that these metrics should not be used as evaluation tools (Ortega, 2019). A closely related literature is open access (OA) which was placed on the map adjacent to the altmetrics and policy clusters.

A large-scale analysis determined that a significant proportion of the literature is freely accessible, and that the open access citation advantage is about 18%, taking into account discipline and age, which was mainly due to specific OA mechanisms (Piwowar et al., 2018). However, more recent science-wide analysis muddies the picture: the advantage of open access is neither consistent across disciplines nor for citations of more recent papers, nor is it more pronounced for the metrics that measure online attention (such as news, blogs, social media) (Taylor, 2026). The number of open documents in our own corpus, which reflects the underlying growth of openness, increases from 20.7% to 68.0% between the two periods (2000-2004 and 2020-2024, see figure 4). The message of the two "impact" literatures - altmetrics and OA - converges and it is this: "impact" is not a single thing. It's a family of different signals, each with distinct behaviors across fields, and no single indicator will do: a pluralism that the responsible-metrics movement begins with.

The computational turn: AI, machine learning and NLP. The breakthroughs examined thus far are based on bibliographic data, that is, who cited whom and which words co-occur. The most significant change in recent times has been the transition from metadata to full text, supported by the new technologies of natural language processing (NLP) and machine learning (ML). This computational shift will alter the fundamental level of analysis: citations can no longer simply be counted; they can be read, interrogated, and asked "why, how, and in what spirit" are they being cited? There are two broad areas of growth that are well documented in the systematic review of the area: semantic analysis of citations and representation and discovery of scientific knowledge (Saeidnia et al., 2025; Safder et al., 2023).

The premise of citation-based evaluation has always been known to be false; citations serve many rhetorical and social functions, and a citation that contradicts a work counts the same as one that builds on it. This long-standing critique is made operational in the computational turn. Decades of small manual research have been synthesized in a single classification that distinguishes between a scientific motivation (rhetoric) and a tactical motivation (social/benefit oriented) (Lyu et al., 2021). NLP allows such distinctions to be measured on scale. Previous supervised studies demonstrated that features like the number of citations and citation location in the text can be automatically used to identify influential citations (Valenzuela et al., 2015).

Later large-scale research categorized citations at the whole field level, and found that framing a paper's citations predicts the number of citations it will receive, and that such changes mirror the development of a field towards agreement (Jurgens et al., 2018). This results in a number of subfields that are then mapped as a coherent program in a narrative review: citation context and content analysis, citation classification, sentiment analysis, summarisation and recommendation, and all based on full-text processing (Saeidnia et al., 2025). These techniques have also been extended from individual research projects to large-scale infrastructures that classify citation contexts as supporting, contrasting, or mentioning cited work (Saeidnia et al., 2025). Altmetric

data pipelines for sentiment analysis also try to extract the valence (or positivity) of online mentions, not just the number of mentions (Hassan et al., 2021).

The second front is dealing with the representation of scientific documents computationally. There are several of pre-trained language models that have been adapted for scientific text, such as SciBERT, trained on a large multidomain collection of scientific articles, which have significantly outperformed general-purpose pre-trained models in downstream scientific NLP tasks. To generate document-level representations instead of token-level representations, SPECTER utilizes the citation graph as a training signal to generate embeddings which can be used for classification and recommendation without fine-tuning for a specific task (Cohan et al., 2020).

Neural topic-modelling methods such as BERTopic, rather than named-entity-recognition methods, use the transformer embeddings and introduce a class-based weighting scheme to generate coherent topics with clear interpretation, which is an alternative to surface co-occurrence mapping by applying semantic analysis (Grootendorst, 2022). Maybe the most impressive example of latent knowledge in the literature is in materials science: unsupervised word embeddings trained on past publications encoded chemical and structural relationships, and that knowledge was sufficiently explicit to be used successfully for predicting functional materials years before any were discovered, indicating that the published corpus contains recoverable knowledge not yet explicit to its authors (Tshitoyan et al., 2019). All of this rests on the fact that there are large, structured scholarly datasets available, such as large collections of computer science metadata that can be used to analyse trends in the field at a scale of thousands of papers or more (Wahle et al., 2022); it also relies on ML approaches that rank and prioritise research-assessment parameters empirically, not by convention (Ahmed et al., 2023). The computational turn isn't just transformative because it allows an analysis of the content of communications to be performed on something other than metadata; it's transformative because it allows the semantics of scholarly communication to be measured. It brings in new dangers that field is just starting to face, though.

The outputs of classifiers are often dependent on training data and model selection that is not easily understood, and a citation-intent label or a sentiment score comes with some level of uncertainty that a raw count does not; adding up the outputs of the classifiers to make evaluation indicators can hide model error as apparent precision. Reproducibility is a special issue, because the field studies (Kapoor & Narayanan, 2023; Wang et al., 2026) are subjected to the same publish-or-perish pressures as scientific analyses based on ML, and also because of the possibility of very subtle data leakage. The methodological norms that it has established for citations (transparency of construction, explicit reporting of assumptions, validation against external criteria) are equally applicable and, indeed, more so with the indicators derived from AI, whose inner workings are less accessible to inspection.

The computational turn, briefly, is one that seeks to expand the scope of what scientometrics can measure, but not at the cost of withholding the interpretive reservations that the remainder of this review has amplified. Research evaluation, policy and the constitutive effects of metrics. The previous subsections have dealt with indicators as tools of measurement. This one treats them as the means of governance, because in contemporary scientometrics, what is at stake is the allocation of jobs, grants and prestige, and its application alters the behaviour of the scientists under measurement.

The key theme of the evaluation literature is that measurements do not simply reflect scientific processes but are rather actively involved in them. It was the reform movement that crystallised in two significant statements. The San Francisco Declaration on Research Assessment (DORA) challenged specifically the use of journal-level measures, especially impact factors, as a proxy for the quality of individual articles, or individual researchers (Gagliardi et al., 2023).

The Leiden Manifesto extended the argument to a number of principles for the responsible use of quantitative indicators: metrics should supplement, never replace, expert judgment; differences in the field should be taken into account during evaluation; indicators should be transparent and verifiable; and the indicators evaluated should have the possibility to check the data. Both statements have a common diagnosis that runs through this review: An indicator without a question to which it responds, and used regardless of its construction and limitations, has no objectivity, but distortion. This position is strengthened by the increasing acceptance of the idea that citations are imperfect, context-dependent proxies with meaning that relies on the motivating factors and discipline (Aksnes et al., 2019; Ioannidis et al., 2016).

Strategic behaviour studies are the best available evidence that metrics are constitutive. Once the indicator is becoming a goal, the persons measured for it change their behavior to maximize this indicator (the phenomenon is known as “gaming”). Italy's “inwardness” (national share of citations) increased dramatically following the national introduction of bibliometric requirements for career advancement in 2011; this resulted in Italy becoming an outlier among similar countries, likely because of the strategic use of self-citations and citation clubs in response to the evaluation system (Baccini et al., 2019).

The case shows that an evaluation system not only measures a research community but, it changes the citation behaviour of a research community. This is not an unintended consequence of having metrics to govern that needs to be fixed, but rather an inherent characteristic of metric-based governance, and as such, a hard limit on the amount of any given indicator that should be relied on without it being compromised by incentives it creates. A more radical approach to the literature on science and technology studies claims that the process of evaluation do not merely represent or distort existing science but is also a process that helps construct fields of science. A study of how US ‘translational science’ came into being reveals that evaluation practices do not only measure the impact of translational science, but shape its emergence, develop representations of it for researchers, funders and policymakers, and contribute to its legitimation as a form of science that is worthy of public investment (Molldrem et al., 2025) - a way of co-production between science and metrics that its authors describe as ‘the evaluative way’.

The strongest version of the claim, whether accepted or not, serves to focus the central lesson: that scientometric indicators are instruments that act upon the science they measure, and thus require an act of science policy to design; and that the responsibility for producing the impacts of scientometric indicators cannot be divorced from the responsibility for the indicators' accuracy. Synthesis and Open Challenge. Three cross-cutting tensions structure the field's agenda. Coverage vs. comparability: each indicator is contingent on a source that can be biased, and open infrastructure can be about curation, but it cannot be about noise; methods need to be explicit about the need for source limits, and need to be robust to them. The terms quantification versus validity: impact is plural, so the challenge is to not reduce it to single scores but to correlate indicators with questions. The computable turn in the discipline: the way to making semantics measurable is making them opaque in the process of automation: the transparent and validated rules of the discipline should apply to the indicators created with AI. In all three, the constitutive knowledge is that the methodological decisions in scientometrics are decisions about the governance of science.

Conclusion

Scientometrics has evolved from a narrow focus on citation indexing to a discipline and approach that influence the way science is funded, evaluated and organised. This review was a synthesis of five methodological layers, bound together by a consistent commitment to evaluation and policy. Two findings stand out. First, the most enduring advances in the field have been not in the introduction of new indicators, but in the improvements of the underlying data infrastructure and greater transparency about the context in which these data are evaluated. New indicators are not the most important advances in the field, however, but improvements in the underlying data infrastructure, and greater candor about the evaluation context surrounding it. Second, the new methodological paradigm, from counting to understanding citations, is the most radical development in a generation but it should be subject to the same standards of transparency, validity, and responsible use that were developed for the simpler methods; indeed, the more powerful the methods, the more important those standards. The accompanying analysis is based on an open and reproducible data source and is presented as evidence of the structure of the field, as well as an example of the transparency that the field increasingly demands of itself.

Conflict of Interest Statement

The authors declare no potential conflicts of interest with respect to the research, authorship, or publication of this article.

Author Contributions

Magzhan Amangeldi and Oleksandr Kuchanskyi contributed equally to the conceptualization, methodology, data collection and processing, bibliometric analysis, interpretation of the findings, preparation of the original manuscript, review, editing, and finalization of the study. Both authors approved the final

version of the manuscript.

Declaration of Use of Generative AI and AI-Assisted Technologies

The authors declare that no artificial intelligence (AI) tools were used in the preparation of this manuscript.

References

- Abramo, G., Cicero, T., & D'Angelo, C. A. (2026). Data source effects in research performance assessment of individuals and institutions: Comparing OpenAlex with established bibliographic databases. *Scientometrics*, 131(5), 3411–3441. <https://doi.org/10.1007/s11192-026-05638-6>
- Ahmed, B., Wang, L., Al-Shamayleh, A. S., Afzal, M. T., Mustafa, G., Alrawagfeh, W., & Akhunzada, A. (2023). Machine learning approach for effective ranking of researcher assessment parameters. *IEEE Access*, 11, 133294–133312. <https://doi.org/10.1109/ACCESS.2023.3336950>
- Aksnes, D. W., Langfeldt, L., & Wouters, P. (2019). Citations, citation indicators, and research quality: An overview of basic concepts and theories. *SAGE Open*, 9(1), 2158244019829575. <https://doi.org/10.1177/2158244019829575>
- Baccini, A., De Nicolao, G., & Petrovich, E. (2019). Citation gaming induced by bibliometric evaluation: A country-level comparative analysis. *PLOS ONE*, 14(9), e0221212. <https://doi.org/10.1371/journal.pone.0221212>
- Bornmann, L., & Haunschild, R. (2017). An empirical look at the Nature Index. *Journal of the Association for Information Science and Technology*, 68(3), 653–659. <https://doi.org/10.1002/asi.23682>
- Cohan, A., Feldman, S., Beltagy, I., Downey, D., & Weld, D. S. (2020). SPECTER: Document-level representation learning using citation-informed transformers. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 2270–2282). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.acl-main.207>
- Culbert, J. H., Hobert, A., Jahn, N., Haupka, N., Schmidt, M., Donner, P., & Mayr, P. (2025). Reference coverage analysis of OpenAlex compared to Web of Science and Scopus. *Scientometrics*, 130(4), 2475–2492. <https://doi.org/10.1007/s11192-025-05293-3>
- Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, 133, 285–296. <https://doi.org/10.1016/j.jbusres.2021.04.070>
- Ellegaard, O. (2026). The method of bibliometric analysis: A critical evaluation. *Scientometrics*, 131, 4175–4196. <https://doi.org/10.1007/s11192-026-05670-6>
- Faghri, A., & Bergman, T. L. (2024). Highly ranked scholars and the influence of countries and regions in research fields, disciplines, and specialties. *Quantitative Science Studies*, 5(2), 464–483. https://doi.org/10.1162/qss_a_00291
- Fortunato, S., Bergstrom, C. T., Börner, K., Evans, J. A., Helbing, D., Milojević, S., Petersen, A. M., Radicchi, F., Sinatra, R., Uzzi, B., Vespignani, A., Waltman, L., Wang, D., & Barabási, A.-L. (2018). Science of science. *Science*, 359(6379), eaao0185. <https://doi.org/10.1126/science.aao0185>
- Gagliardi, A. R., Chen, R. H., Boury, H., Albert, M., Chow, J., DaCosta, R. S., Deyell, R., Finelli, A., Keating, A., Légaré, F., McGowan, J., McLeod, R., Moore, A., Parulekar, W., Saginur, R., Straus, S. E., & Protze, S. (2023). DORA-compliant measures of research quality and impact to assess the performance of researchers in biomedical institutions: Review of published research, international best practice and Delphi survey. *PLOS ONE*, 18(5), e0270616. <https://doi.org/10.1371/journal.pone.0270616>
- Garfield, E. (1955). Citation indexes for science: A new dimension in documentation through association of ideas. *Science*, 122(3159), 108–111. <https://doi.org/10.1126/science.122.3159.108>
- Genc, H. N., & Kocak, N. (2024). Bibliometric analysis of studies on artificial intelligence in science education with VOSviewer. *Journal of Education in Science, Environment and Health*, 10(4), 183–195. <https://doi.org/10.55549/jeseh.756>
- González, P., Fors, M., & Torres, A. (2025). Altmetrics in the evaluation of scholarly impact: A systematic and critical literature review. *Frontiers in Research Metrics and Analytics*, 10, 1693304. <https://doi.org/10.3389/frma.2025.1693304>

- Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv*. <https://doi.org/10.48550/arXiv.2203.05794>
- Hassan, S. U., Saleem, A., Soroya, S. H., Safder, I., Iqbal, S., Jamil, S., Bukhari, F., Aljohani, N. R., & Nawaz, R. (2021). Sentiment analysis of tweets through Altmetrics: A machine learning approach. *Journal of Information Science*, 47(6), 712–726. <https://doi.org/10.1177/0165551520930917>
- Haunschild, R., & Bornmann, L. (2024). The use of OpenAlex to produce meaningful bibliometric global overlay maps of science on the individual, institutional, and national levels. *PLOS ONE*, 19(12), e0308041. <https://doi.org/10.1371/journal.pone.0308041>
- Hendricks, G., Tkaczyk, D., Lin, J., & Feeney, P. (2020). Crossref: The sustainable source of community-owned scholarly metadata. *Quantitative Science Studies*, 1(1), 414–427. https://doi.org/10.1162/qss_a_00022
- Hicks, D., Wouters, P., Waltman, L., De Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature*, 520(7548), 429–431. <https://doi.org/10.1038/520429a>
- Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, 102(46), 16569–16572. <https://doi.org/10.1073/pnas.0507655102>
- Hood, W. W., & Wilson, C. S. (2001). The literature of bibliometrics, scientometrics, and informetrics. *Scientometrics*, 52(2), 291–314. <https://doi.org/10.1023/A:1017919924342>
- Hook, D. W., Porter, S. J., & Herzog, C. (2018). Dimensions: Building context for search and evaluation. *Frontiers in Research Metrics and Analytics*, 3, 23. <https://doi.org/10.3389/frma.2018.00023>
- Ioannidis, J. P. A., Boyack, K. W., & Wouters, P. F. (2016). Citation metrics: A primer on how (not) to normalize. *PLOS Biology*, 14(9), e1002542. <https://doi.org/10.1371/journal.pbio.1002542>
- Jurgens, D., Kumar, S., Hoover, R., McFarland, D., & Jurafsky, D. (2018). Measuring the evolution of a scientific field through citation frames. *Transactions of the Association for Computational Linguistics*, 6, 391–406. https://doi.org/10.1162/tacl_a_00028
- Kapoor, S., & Narayanan, A. (2023). Leakage and the reproducibility crisis in machine-learning-based science. *Patterns*, 4(9), 100804. <https://doi.org/10.1016/j.patter.2023.100804>
- Klavans, R., & Boyack, K. W. (2017). Which type of citation analysis generates the most accurate taxonomy of scientific and technical knowledge? *Journal of the Association for Information Science and Technology*, 68(4), 984–998. <https://doi.org/10.1002/asi.23734>
- Leydesdorff, L., Wouters, P., & Bornmann, L. (2016). Professional and citizen bibliometrics: Complementarities and ambivalences in the development and use of indicators—A state-of-the-art report. *Scientometrics*, 109(3), 2129–2150. <https://doi.org/10.1007/s11192-016-2150-8>
- Lu, X., Li, L., & Shen, Z. (2026). Revisiting the field normalization approaches and practices. *Scientometrics*. <https://doi.org/10.1007/s11192-026-05676-0>
- Lyu, D., Ruan, X., Xie, J., & Cheng, Y. (2021). The classification of citing motivations: A meta-synthesis. *Scientometrics*, 126(4), 3243–3264. <https://doi.org/10.1007/s11192-021-03908-z>
- Martín-Martín, A., Costas, R., van Leeuwen, T., & Delgado López-Cózar, E. (2018). Evidence of open access of scientific publications in Google Scholar: A large-scale analysis. *Journal of Informetrics*, 12(3), 819–841. <https://doi.org/10.1016/j.joi.2018.06.012>
- Molldrem, S., Moses, J., Smith, E., Tumilty, E., & Farroni, J. (2025). Metrics producing science: The evaluative way in the emergence of US translational science. *Science & Technology Studies*. <https://doi.org/10.23987/sts.145702>
- Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics*, 106(1), 213–228. <https://doi.org/10.1007/s11192-015-1765-5>
- Moral-Muñoz, J. A., Herrera-Viedma, E., Santisteban-Espejo, A., & Cobo, M. J. (2020). Software tools for conducting bibliometric analysis in science: An up-to-date review. *El Profesional de la Información*, 29(1), e290103. <https://doi.org/10.3145/epi.2020.ene.03>
- Omer, A. A. A., Issa, I. I. M., Abuker, Y. Y. A., & Asad, S. (2025). Evaluating the use of VOSviewer in bibliometric and science-mapping studies: Trends, current state and future directions. *International Journal of Current Research in Science, Engineering & Technology*, 8(4), 483–497. <https://doi.org/10.30967/IJCRSET/Altyeb-Ali-Abaker-Omer/208>

- Ortega, J. L. (2019). Exploratory analysis of Publons metrics and their relationship with bibliometric and altmetric impact. *Aslib Journal of Information Management*, 71(1), 124–136. <https://doi.org/10.1108/AJIM-06-2018-0153>
- Piwowar, H., Priem, J., Larivière, V., Alperin, J. P., Matthias, L., Norlander, B., Farley, A., West, J., & Haustein, S. (2018). The state of OA: A large-scale analysis of the prevalence and impact of Open Access articles. *PeerJ*, 6, e4375. <https://doi.org/10.7717/peerj.4375>
- Radhakrishnan, S., Erbis, S., Isaacs, J. A., & Kamarthi, S. (2017). Novel keyword co-occurrence network-based methods to foster systematic reviews of scientific literature. *PLOS ONE*, 12(3), e0172778. <https://doi.org/10.1371/journal.pone.0172778>
- Saeidnia, H. R., Hosseini, E., Abdoli, S., & Ausloos, M. (2025). Unleashing the power of AI: A systematic review of cutting-edge techniques in AI-enhanced scientometrics, webometrics and bibliometrics. *Library Hi Tech*, 43(4–5), 1292–1314. <https://doi.org/10.1108/LHT-10-2023-0514>
- Safder, I., Ali, M., Aljohani, N. R., Nawaz, R., & Hassan, S. U. (2023). Neural machine translation for in-text citation classification. *Journal of the Association for Information Science and Technology*, 74(10), 1229–1240. <https://doi.org/10.1002/asi.24817>
- Sugimoto, C. R., Work, S., Larivière, V., & Haustein, S. (2017). Scholarly use of social media and altmetrics: A review of the literature. *Journal of the Association for Information Science and Technology*, 68(9), 2037–2062. <https://doi.org/10.1002/asi.23833>
- Taylor, M. (2026). Evaluating open access advantages for citations and altmetrics (2011–2021): A dynamic and evolving relationship. *Quantitative Science Studies*, 7, 643–670. <https://doi.org/10.1162/QSS.a.470>
- Tshitoyan, V., Dagdelen, J., Weston, L., Dunn, A., Rong, Z., Kononova, O., Persson, K. A., Ceder, G., & Jain, A. (2019). Unsupervised word embeddings capture latent knowledge from materials science literature. *Nature*, 571(7763), 95–98. <https://doi.org/10.1038/s41586-019-1335-8>
- Valenzuela, M., Ha, V., & Etzioni, O. (2015). Identifying meaningful citations. In *AAAI Workshop: Scholarly Big Data* (Vol. 15, p. 13).
- van Eck, N. J., & Waltman, L. (2017). Citation-based clustering of publications using CitNetExplorer and VOSviewer. *Scientometrics*, 111(2), 1053–1070. <https://doi.org/10.1007/s11192-017-2300-7>
- Visser, M., van Eck, N. J., & Waltman, L. (2021). Large-scale comparison of bibliographic data sources: Scopus, Web of Science, Dimensions, Crossref, and Microsoft Academic. *Quantitative Science Studies*, 2(1), 20–41. https://doi.org/10.1162/qss_a_00112
- Wahle, J. P., Ruas, T., Mohammad, S., & Gipp, B. (2022). D3: A massive dataset of scholarly metadata for analyzing the state of computer science research. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference* (pp. 2642–2651). European Language Resources Association. <https://doi.org/10.63317/3hvgsmuea6o7>
- Waltman, L., & van Eck, N. J. (2019). Field normalization of scientometric indicators. In W. Glänzel, H. F. Moed, U. Schmoch, & M. Thelwall (Eds.), *Springer handbook of science and technology indicators* (pp. 281–300). Springer International Publishing. https://doi.org/10.1007/978-3-030-02511-3_11
- Wang, X., Gu, P., & Shehzadi, K. (2026). Global open science research in the digital & intelligent era: Advances, focuses and prospects. *Scientometrics*. 10.1007/s11192-026-05597-y
- Weimer, V., Heck, T., van Leeuwen, T., & Rittberger, M. (2023). The quantification of open scholarship—A mapping review. *Quantitative Science Studies*, 4(3), 650–670. https://doi.org/10.1162/qss_a_00266
- Zhang, L., Cao, Z., Shang, Y., Sivertsen, G., & Huang, Y. (2024). Missing institutions in OpenAlex: Possible reasons, implications, and solutions. *Scientometrics*, 129(10), 5869–5891. <https://doi.org/10.1007/s11192-023-04923-y>

Information about authors

Magzhan Amangeldi – Master's degree, Doctoral Student, Astana IT University, Astana, Republic of Kazakhstan; e-mail: 255242@astanait.edu.kz; ORCID iD: 0009-0002-7979-0502.

Oleksandr Kuchanskyi – Doctor of Technical Sciences, Professor, School of Artificial Intelligence and Data Science, Astana IT University, Astana, Republic of Kazakhstan; National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Kyiv, Ukraine; e-mail: kuczanski@gmail.com; ORCID iD: 0000-0003-1277-8031 (*corresponding author*).

Received: 10.07.2026

Revised: 28.08.2026

Accepted: 05.09.2026

Published: 30.09.2026

Type of Article: Research Article

¹Kamilya Arinova*, ²Yerkin Nurmakhanov

¹ Abai Kazakh National Pedagogical University, Almaty, Republic of Kazakhstan

²Karaganda National Research University named after Academician E. A. Buketov, Karaganda, Republic of Kazakhstan

DIGITAL EDUCATIONAL ENVIRONMENT AS A RESOURCE FOR DEVELOPING DIGITAL COMPETENCIES OF FUTURE SPECIAL EDUCATORS: CHALLENGES AND PROSPECTS IN THE REPUBLIC OF KAZAKHSTAN

Abstract. This article examines the perceptions of future special educators regarding the development of digital competencies in the context of the evolving digital educational environment in the Republic of Kazakhstan, focusing on challenges and prospects. The study was guided by four research questions and utilized a descriptive survey research design. The study population consisted of 239 future special educators from three leading higher education institutions in the Republic of Kazakhstan – Abai Kazakh National Pedagogical University, Al-Farabi Kazakh National University and L. N. Gumilyov Eurasian National University. Data were collected between January and March 2026. Given that the population size was manageable, complete enumeration was used; no sampling technique was applied. A 40-item author-developed questionnaire titled “Challenges and Prospects of Forming Digital Education Environments Questionnaire” (CPFDEER) was used. Face and content validity was ensured by three experts in Educational Measurement and Evaluation. Internal consistency was assessed using Cronbach's alpha (Cluster 1 = 0.81, Cluster 2 = 0.82, Cluster 3 = 0.81, Cluster 4 = 0.83, overall $\alpha = 0.82$). Data were analyzed using mean and standard deviation with SPSS version 26.0. The results showed that respondents perceived that digital learning environments support quality training for future special educators. The findings highlight perceived challenges such as technical issues, limited access to technology, and inadequate training. This study provides insights into respondents' perceptions of how digital learning environments support special education.

Keywords: digital education environments, educator training, special education, special educators.

Introduction

Globally, education is part and parcel of man because of the role it plays in the cognitive, affective and psychomotor development of man and his environment in any given community. Education is, therefore, an instrument through which people attain all-around success and become active members of society. The transformation of the education sector due to the advent of information and communication technology (ICT) is creating new opportunities for effective teaching and learning, as well as for collaboration, marking an essential shift from traditional teaching methods to digital approaches (Bond et al., 2020; Zhao et al. 2021). Similarly, the digital education environment emerged due to the integration of ICT and modern technologies in distance learning (Yessenbekova, et al., 2025). Consequently, future special educators who provide instruction to learners with special needs need to pursue specialized training to acquire the relevant skills and ICT competencies to use digital educational environments (Basilotta-Gomez-Pablos et al., 2022; Falloon, 2020). This makes the formation of digital environments for training special educators a necessity.

In simple terms, a digital learning environment is an online platform that provides an interactive, comprehensive learning experience. According to Fernández-Batanero et al. (2022), the digital education environment encompasses various resources and digital tools, such as online libraries, multimedia resources, learning management systems, and virtual classrooms. However, all these digital tools and resources provide learners with maximum access to a wealth of information, help foster good communication and collaboration among educators and learners, and also offer personalized learning experiences tailored to the needs and abilities of individuals (Ferrer et al., 2022; Kerimbayev et al., 2023). Besides, the digital education environment tends to prepare learners for the demands of the digital age, promote the quality of teaching and learning, and increase learners' academic engagement and motivation (Zhao et al., 2023; Claro et al., 2024). However, several problems impede the formation of a digital education environment for training special

educators. One of the key problems appears to be a lack of infrastructure, such as poor internet connectivity, inadequate computer hardware and digital resource, a challenge well documented during the COVID-19 pandemic and in Kazakhstan context (OECD, 2021; Yakavets et al., 2023).

This problem affects the growth and development of the digital education environment, making it very difficult for future special educators to participate effectively in virtual learning activities and access online resources. In addition, it appears that the lack of infrastructure is driven by poor funding, making it challenging to maintain existing digital equipment and procure new equipment. It appears that inadequate educator training constrains the digital education environment.

According to Guillén-Gámez and Mayorga-Fernandez (2020), future special educators need to be trained to use digital technologies in teaching, as most are ill-equipped with the relevant knowledge and skills. Thus, inadequacy in skills and knowledge regarding how to integrate digital tools into their teaching practices impede the effectiveness of the digital learning environment and the future special educators' ability to provide quality education to learners with special educational needs (Dias-Trindade et al., 2020). In addition, inadequate maintenance and technical support appear to have contributed to the problem of inadequate training. This makes it Herculean for future special educators to troubleshoot technical problems and also with their teaching activities (Dang et al, 2024; De Obesso et al., 2023).

Differently stated, inequality and the digital divide in accessing digital technologies tend to affect the formation of digital education environments, particularly for training future special educators (Claro et al., 2024). Digital divide refers to differences in access to technology among individuals, regions, and institutions. This divide appears to be common in developing societies. In the context of the Republic of Kazakhstan, which is implementing inclusive education reforms under the state program for Education Development 2020-2025, this challenge remains relevant (Yessenbekova et al., 2025). Consequently, Ainscow (2020) posited that educators working in these societies are left behind, as it is more challenging to access online resources and effectively participate in digital education environments. Another challenge, according to Fernandez-Batanero et al. (2022), that needs to be addressed is inadequate accessibility and inclusivity. This is because digital education environments may often neglect students' needs, especially those with special educational needs.

Despite these factors that seem to impede digital education environments, the prospects for training future special educators remain promising. As Cobero-Almenara et al. (2022) stated, future special educators have access to technologies, and innovative teaching methods through digital educational environments. Through this, the skills and competencies of future special educators in teaching learners with special educational needs are enhanced (Dang et al. 2024; Gallego Joya et al., 2025). Digital education environments create opportunities for collaboration, networking, and professional development among special educators, which go a long way toward enhancing their job retention and satisfaction. Besides, there are very few studies that address the problems and prospects of forming digital education environments for training special educators in Central Asian context, and filling this knowledge is germane (Basilotta-Gomez-Pablos et al., 2022). Based on the foregoing, it becomes prudent to examine the challenges and prospects of forming digital education environments for training future special educators in the Republic of Kazakhstan.

Statement of the Problem

In many developed and developing societies including the Republic of Kazakhstan, there appear to be many constraints on the development of digital education environments for training future special educators, such as limited access to digital resources, inadequate infrastructure, and insufficient educator training (OECD, 2021; Yakavets et al., 2023). Consequently, it becomes a challenging task for future special educators to acquire quality teaching, particularly in this period of rapid technological growth and development. The current situation is that the digital education environment for future special educators appears to be characterized by traditional teaching and learning methods that often neglect the diverse abilities and needs of learners with special needs, leading to limited accessibility and inclusive education (Ainscow, 2020; Fernandez- Batanero et al., 2022).

If these challenges confronting the formation of digital environments for training future special educators are not addressed, the majority of learners, particularly those with special educational needs, will be left behind. To the best of knowledge, the best practices for the formation of digital environments to address the diverse needs of future special educators in Kazakhstan have not yet been identified. Part of the problem stems from the paucity of empirical studies on digital education environments for future special educators in this context. This

leaves educators and educational policymakers with limited guidance on how to use digital technologies in special education programs (Zhao et al., 2021; Claro et al., 2024). Hence, it becomes needful to examine the perceptions regarding challenges and prospect of the digital education environment for training future special educators in Kazakhstan.

Research Aims

The general aim of this research is to examine the perceptions regarding challenges and prospects of forming digital education environments for training future special educators. Specifically, the study sought to determine:

- 1) The ways of forming digital education environments to enhance the training of future special educators
- 2) The problems associated with the formation of digital education environments for future special educators.
- 3) How can the problems confronting the formation of digital education environments for future special educators be reduced?
- 4) How the prospects of digital education environments can be used as a tool for improving the training of future special educators.

Research Questions

For the purpose of addressing the above research aims, the following questions were raised:

- 1) In what ways can digital education environments be formed to support the training of future special educators?
- 2) What are the problems confronting the formation of digital education environments for future special educators?
- 3) What measures can be adopted to reduce the problems confronting the formation of digital education for future special educators?
- 4) In what ways can the prospects of a digital education environment be used for improving the training of future special educators?

Significance of the Study

The findings of this research would be of immense benefit to educational stakeholders, including students with special educational needs, special educators, educational institutions, educational policymakers, curriculum planners, and future researchers, when this work is completed and published in journals.

The findings of this research will be of immense benefit to students with special educational needs, as they will be informed about the problems and prospects of forming a digital education environment. This will go a long way towards enhancing their access to quality and inclusive education system. Moreover, the findings of the study will provide learners with special educational needs with personalized, flexible educational experiences that cater to their needs and abilities. This is achievable when the work is published in journals and proceedings that students can access.

The result of this research will be of immense benefit to special educators by providing insights into effective implementation and by designing digital educational environments that address the diverse needs and abilities of learners with special educational needs. This will go a long way towards enhancing their skills and competencies required in teaching learners with special educational needs, particularly when digital resources are geared towards meeting the diverse needs of students. Special educators will be well-equipped with this finding by attending and participating in conferences, seminars, and workshops where papers on digital education environments are presented and discussed.

Educational institutions will also benefit from the findings of this research, as they will be guided in developing and implementing digital educational environments with the sole aim of supporting future special educators' training. This will enhance the academic achievement of learners with special educational needs and promoting job satisfaction among special educators. This will be achieved through the publication of journals and consultation with the educational administrators.

The findings of this research will be of immense help to educational policymakers, as they will be better informed regarding the development of suitable guidelines and policies for the formation of digital education environments in special education. By doing so, access to quality education for students with special educational needs will be improved. This will be achieved through journal publications and consultations with educational policymakers worldwide.

For curriculum planners, the findings of this study will help them consolidate effective teaching in

digital education environments for pre-service special educators. This is achievable when this study is published online in a journal for easy consultation by curriculum planners in the educational system.

Finally, future researchers will find the outcome of this study interesting. This is because the findings of this research will add to the already existing body of knowledge on the formation of digital learning environments. This study will serve as a baseline for other researchers to build on. Hence, researchers conducting studies on special education and digital learning environments will find the study beneficial.

Literature review

Concept of Digital Education Environment

Digital education environment as a concept has become a matter of concern to so many educational stakeholders, such as curriculum planners, educational policy makers, government, school authorities, researchers, among others, because of the advent of digital technologies in education, as well as the need to jettison traditional approaches and embrace innovative approaches to teaching and learning. As noted by Ferrer et al. (2022), the digital education environment is a virtual platform that provides students with comprehensive, interactive learning experiences.

The digital education environment includes various resources and tools, such as virtual classrooms, multimedia content, learning management systems, and digital libraries (Yessenbekova et al., 2025; Kerimbayev et al., 2023). In the special education context, digital environment in education is an online tool that provides future special educators with the relevant resources and tools needed for quality education, including assistive technologies for inclusion (Fernandez-Batanero et al., 2022; Ainscow, 2020).

According to Falloon (2020) and Basilotta-Gomez-Pablos et al. (2022), digital educational environments support learners with special educational needs in accessing quality teaching and learning within an inclusive education system through teacher's digital competence. Some attributes of a digital education environment include adaptability, flexibility, and accessibility, which allow learners to learn at their convenience and time (Guillén-Gámez & Mayorga-Fernández, 2020; Fallon, 2020). Digital education environments serve as a vehicle for networking, professional development, and collaboration among future special educators (Bond et al., 2020).

By way of synthesis, one can infer that a digital education environment is a critical attribute of a special education program, hinged on providing learners with special educational needs access to quality teaching and learning, as well as all-inclusive education. In this study, the digital education environment is regarded as a virtual platform that supports future special educators with resources and tools to promote quality learning among students with special educational needs, fostering equity, inclusion, and diversity.

The formation of digital education environments to enhance the training of future special educators

Forming a digital education environment is a process that involves integrating technology, collaboration, and pedagogy to enhance the training of exceptional educators. According to Zhao et al. (2021), digital education environments foster access to technologies, innovative teaching approaches, and resources, particularly among future special educators. Online and blended learning are among the approaches to the formation of digital education environments (Bond et al., 2020). As stated by Bravo-Agapito et al. (2021), the diverse needs of future special educators are met through online and blended learning, and student-centered approaches using modern technologies. Another way to create digital education environments is to incorporate modern technologies such as artificial intelligence, virtual reality, and augmented reality into special educators' training. As opined by Claro et al. (2024), these digital technologies can provide interactive and immersive learning experiences. In-service training for educators should be carefully channeled to incorporate digital technologies and aligned with DigCompEdu frameworks (Cabero-Almenara et al., 2020; Vuorikari et al., 2022; Dias-Trindade et al., 2020).

By way of recapitulation, it was deduced that the formation of a digital education environment to enhance future special educators' training requires a systematic approach that integrates digital technologies, collaboration, and pedagogy. When these approaches are effectively utilized, special educators develop the required competencies and skills to provide quality instruction to learners with special educational needs.

The challenges confronting the formation of digital education environments for future special educators

There appear to be several problems affecting the formation of digital education environments, which also impede the effective use of digital technologies in educating learners with special needs. As OECD (2021)

rightly stated, one of the key problems is inadequate infrastructure, including digital resources, computer hardware, and internet connectivity. There seems to be a problem of inclusivity and accessibility in the formation of digital education environments (Fernández-Batanero et al., 2022).

Furthermore, another constraint on the formation of a digital education environment for special educators is inadequate technical support and maintenance and perceptions of educators' competence (Dang et al., 2024; De Obesso et al., 2023). In the absence of technical support, future special educators may find it extremely difficult to use digital technologies effectively. From another perspective, Zhao et al. (2021) noted that inadequate in-service training for future special educators is another key factor impeding the development of digital environments, as it limits their ability to integrate effective digital technologies into their pedagogical practices.

To sum up, the problems confronting the formation of digital education environments for the training of future special educators appear to be complex and therefore require a comprehensive solution. When these problems are clearly understood, educators and educational policymakers will have the opportunity to work together to create effective, inclusive, and accessible digital education environments that support the diverse learning needs of learners with special educational needs.

Strategies for reducing the problems confronting the formation of digital education environments for the training of future special educators

Salvaging the problems affecting the formation of digital education environments for the training of future special educators requires a comprehensive approach that includes investment, collaboration, and innovation. According to Gallego et al. (2025), providing continuous professional development and targeted training for future special educators is considered one of the key strategies. Investment in infrastructure such as digital resources, computer, hardware and reliable internet connectivity is also crucial (Kerimbayev et al. 2023). In addition, fostering stronger collaboration among educational policymakers, educators, and digital technology providers is essential. Such collaborative efforts can facilitate the development of inclusive and accessible digital education environments capable of addressing existing challenges. This will help address the needs of exceptional educators, including those of learners with special needs. Apart from that, educators could also engage in effective collaboration among themselves, as it creates room for them to share the best resources and practices, reduce excessive workload, and enhance the effectiveness of the digital learning environment (OECD, 2021). Evidence from Kazakhstan pedagogical universities shows that an integrated approach combining technology-enhance pedagogy significantly improved digital competence (Dang et al., 2024; De Obesso et al. 2023). Conclusively, reducing the problems associated with the digital education environment for the training of future special educators requires comprehensive strategies, investment and collaboration.

The prospects of digital education environments as a tool for improving the training of future special educators

The prospect of digital education environments offers many ways to improve future special educators' training and the academic performance of learners with special educational needs. As Zhao et al. (2021) noted, future special educators could have maximum access to various digital technologies, resources, and innovative teaching approaches through the integration of digital education environments. Digital competence is a strong predictor of perceived learning value and inclusion in Kazakhstan context (Dang et al 2024, Yessenbekova et al., 2025). Learning analytics serve as a platform through which future special educators can identify areas where learners require greater support (Bond et al. 2020). The second point to note is to effectively utilize digital education environments to enhance networking and collaboration among special educators. To further support this point, digital education environments enable students with special needs to gain greater access to assistive and adaptive technologies, as well as personalized learning plans. Moreover, parent-educators engagement and effective communication can be improved through digital education environments (Bravo-Agapito et al., 2021). To summarize, it was deduced that the prospect of digital education environments offers different opportunities to improve future special educators' training and the academic achievement of learners with special educational needs. When these prospects are leveraged, educators and other educators create digital education environments that are highly effective and innovative, helping address the diverse challenges affecting students with special needs in the 21st century.

Gap Identification

The literature reviewed so far has revealed an apparent gap in the formation of digital education environments for future special educators' training in Kazakhstan context. Although studies reviewed showed the merits of digital education environments, there appears to be a lack of empirical evidence on the problems and prospects of integrating them specifically for future special educators in Central Asian context. While existing studies focuses largely on general online learning (Bond et al., 2020; Ferrer et al., 2022), studies specifically on digital competencies are limited (Fernandez-Batanero et al., 2022, Cabero-Almenara et al., 2022; Basilotta-Gomez-Pablos et al., 2022).

Theoretical Framework

This study adopted two theories: Technological Pedagogical Content Knowledge (TPACK) and Self-Determination Theory (SDT) to advance its argument.

Technological Pedagogical Content Knowledge (TPACK)

The theory states that effective teaching in a digital education environment depends on integrating three aspects of knowledge: content knowledge, pedagogical knowledge, and technological knowledge. The affinity among these three domains of knowledge creates an intersection of additional knowledge, including the central TPACK, as well as pedagogical content knowledge, technological pedagogical knowledge, and technological content knowledge. In relation to the present study, TPACK theory informs the investigation of future special educators' ability to effectively integrate emerging technologies into their teaching practices to enhance learners' academic outcomes. Thus, TPACK theory underscores the relevance of special educators possessing Technological Pedagogical Content Knowledge (TPACK), as it enables them to utilize digital technologies and resources in their pedagogical practices, which will go a long way toward enhancing the academic performance of learners with special needs. Thus, when content, digital technologies, and pedagogy are integrated into digital education environments, special educators gain many opportunities to create an engaging, all-inclusive academic environment that addresses the academic problems of students with special needs.

Self-Determination Theory (SDT)

Self-Determination theory posits three innate psychological needs: relatedness, autonomy, and competence. These three innate psychological needs underpin intrinsic motivation and optimal learning. In this research, self-determination theory was used to argue that personalized learning paths and flexible pacing often support autonomy in digital education environments. It is also important to affirm that competence is enhanced through constructive feedback, while clear, related objectives are nurtured through collaborative discussions and group projects. When these needs are satisfied, academic engagement, academic performance and persistence among learners with special education needs and future special educators are enhanced. This theory, therefore, serves as a tool for designing interventions that foster academic performance, motivation, and engagement for special educators, particularly in a digital education environment.

Be that as it may, the two theories reviewed so far complement each other in the context of the present study. While TPACK serves as a framework for effectively integrating content, digital technology and pedagogy in the formation of a digital learning environment, SDT provides platforms for addressing the psychological needs that impede learning and academic motivation. Through the combination of these theories, this research is geared towards investigating how digital education environments could be structured to support future special educators' training and to improve the academic performance of learners with special education needs.

Methods

This subheading provides a clear description of the methods used in conducting the present study, presented as follows: research design, study population, sample and sampling technique, instrument for data collection, instrument validation, instrument reliability, method of data collection, and method of data analysis.

This research employed a descriptive survey design. A descriptive survey research design was considered suitable for the present study because it is an opinion-type study that sought to examine the perceptions of special educators regarding the challenges and prospects of forming digital education environments through the use of questionnaires.

The study population consisted of 239 future special educators studying at three leading higher education institutions in the Republic of Kazakhstan – Abai Kazakh National Pedagogical University, Al-

Farabi Kazakh National University and L.N. Gumilyov Eurasian National University at the time of the study (2025/2026 academic session).

Given that the population size was manageable, no sampling technique was applied, and the entire population was included in the study (Complete enumeration). Eligibility criteria were: final-year bachelor's students in special education programs, willingness to participate, and availability during data collection. Recruitment period was January to March 2026 (Table 1).

Table 1
Demographic characteristics of respondents

VARIABLE	CATEGORY	FREQUENCY	PERCENTAGE (%)
GENDER	Female	181	75.7%
	Male	58	24.3%
AGE	20-22 years	142	59.4%
	23-25 years	81	33.9%
	26 years and above	16	6.7%
UNIVERSITY	Abai Kazakh National Pedagogical University	92	38.5%
	Al-Farabi Kazakh National University	78	32.6%
	L.N. Gumilyov Eurasian National University	69	28.9%
LEVEL	Final-year Bachelor (4th year special education)	239	100.0%
TOTAL		239	100.0%

This research was conducted using a questionnaire designed by the researchers. The instrument is titled “Challenges and prospects of forming digital education environments questionnaire” (CPFDEEQ). The instrument consists of two sections: A and B. Section A sought information on the demographic variables of study participants, while the information needed to answer the four research questions was addressed in Section B. Section B was therefore divided into four clusters.

The first cluster sought the opinions of study participants on ways to form digital education environments to enhance the training of future special educators. The second cluster sought the study participants' opinions on the problems associated with the formation of digital education environments for future special educators. The third cluster sought the study participants' opinions on how the problems confronting the formation of digital education environments for future special educators can be reduced.

In contrast, the fourth cluster sought the opinion of study participants on how the prospects of digital education environments can be perceived as a tool to improve the training of future special educators. The instrument has a total of 40 items, based on a four-point scale: Strongly Agree (SA)-4, Agree (A)-3, Disagree (D)-2, and Strongly Disagree (SD)-1. The decision threshold is mean > 2.50 considered as agreed.

To ensure that the instrument was validated correctly, three copies of the instrument, along with the study's topic, objectives, and research questions, were submitted to three experts in the educational measurement and evaluation unit at Abai Kazakh National Pedagogical University. The corrections and inputs from the experts led to certain modifications, such as the addition of some items and the correction of typographic errors, before the final production of the instrument.

The reliability of the instrument was determined through a pilot test of 20 future special educators in another public university, which is outside the study location. The scores obtained were analyzed using Cronbach's alpha. The reliability coefficients were 0.81 for the ways of forming digital education environments to enhance the training of special educators, 0.82 for the problems associated with the formation of digital education environments for special educators, and 0.81 for how the problems confronting the formation of digital education environments for special educators can be reduced.

In contrast, 0.83 was obtained for items addressing how the prospects of digital education environments can be used as a tool to improve the training of special educators, with an overall reliability coefficient of 0.82. Thus, the instrument was considered reliable for collecting data for the study. According to Taber (2018), the reliability coefficients of 0.81 and above is acceptable.

Ethical approval was obtained from the Research Ethics Committee of Abai Kazakh National Pedagogical University, Approval No. ABAI-REC-2026/02-14/SEP dated 15th January 2026. Informed

consent was obtained from all participants. Participation was voluntary, confidentiality and anonymity were assured, no personal identifiers were collected, and participants had the right to withdraw at any time without consequences

The researchers administered the instrument to the study participants with the help of three research assistants who were future special educators in the institution. The instrument was administered to future special educators during their workshops and seminars between January and March 2026 using the Direct Delivery Method (DDM). All the 239 copies of the questionnaire were carefully retrieved and used for analysis (100% return case).

In this research, the mean and standard deviation were used as statistical tools for data analysis with SPSS version 26.0. The mean was used to answer the questionnaire items in accordance with the research questions. In contrast, the standard deviation was used to test for the homogeneity or otherwise of the mean ratings of study participants.

For decision-making, items with mean scores of 2.5 and above were considered agreed upon by the study participants. It is important to note that these mean rating reflect respondents' perceptions and do not prove that digital environments enhance instruction or student performance, as this is a descriptive survey.

Results

Research Question 1: In what ways can digital education environments be formed to support the training of special educators?

Data generated in Table 2 reveal that providing technical support for educators was rated 3.50, which was followed by providing flexibility in learning experiences and encouraging educators to use digital tools in teaching, with mean scores of 3.45 and 3.40, respectively, indicating that these are significant ways digital education environments can be formed to support the training of future special educators.

Nevertheless, the study participants accepted all the questionnaire items, as these items received mean ratings above 2.50, with a grand mean score of 3.28 and a standard deviation of 0.64. This implies that the respondents' perceived that the formation of digital education environments supports the training of future special educators.

Table 2

Mean rating and standard deviation of study participants on the formation of digital education environments to support quality training for future special educators

S/N	The ways digital education environments can be formed to support the training of special educators include:	Mean	SD	Decision
1.	Enhancing access to training resources	3.21	0.62	Agreed
2.	Providing flexibility in learning experiences	3.45	0.58	Agreed
3.	Facilitating collaboration among special educators	3.01	0.71	Agreed
4.	Providing opportunities for the development of technical skills	3.35	0.65	Agreed
5.	Supporting the diverse needs of special educators	3.12	0.68	Agreed
6.	Enhancing the quality of training for special educators	3.28	0.63	Agreed
7.	Encouraging educators to use digital tools in teaching	3.40	0.60	Agreed
8.	Increasing access to professional development opportunities	3.22	0.66	Agreed
9.	Providing technical support for educators	3.50	0.55	Agreed
10.	Improving the outcome of learners for special educators	3.18	0.69	Agreed
	Grand mean for research question one	3.28	0.64	Agreed

Research Question 2: What are the problems confronting the formation of digital education environments for special educators

In Table 2, inadequate training regarding the use of digital tools, poor internet connectivity, and high-cost technology were regarded as the main problems confronting the formation of digital education environments for future special educators, with mean scores of 3.42, 3.38, and 3.32, respectively.

Furthermore, all the questionnaire items were rated as agreed by the study participants, with a cluster mean score of 3.21 and a standard deviation of 0.67. This indicates that the respondents perceive items in Table 3 as the key problems affecting the formation of digital education environments for future special

educators.

Table 3

Mean rating and standard deviation of study participants on the problems confronting the formation of digital education environments for future special educators

S/N	The problems affecting the formation of digital education environments include:	Mean	SD	Decision
1.	Poor access to technology	3.30	0.64	Agreed
2.	Inadequate training regarding the use of digital tools	3.42	0.59	Agreed
3.	Technical issues	3.20	0.67	Agreed
4.	Poor internet connectivity	3.38	0.61	Agreed
5.	Negative attitude of educators towards the use of digital tools	3.05	0.72	Agreed
6.	Lack of skills	3.15	0.70	Agreed
7.	Digital education environments are not user-friendly	2.90	0.75	Agreed
8.	Inadequate technical report	3.25	0.66	Agreed
9.	High-cost technology	3.32	0.63	Agreed
10.	Lack of institutional support	3.10	0.71	Agreed
11.	Grand mean for research question two	3.21	0.67	Agreed

Research Question 3: What measures can be adopted to reduce the problems confronting the formation of digital education?

The findings in Table 4 reveal that provision of maximum technical support (3.45), provision of quality training on digital tools (3.40), Provision of institutional support (3.35), improving internet connectivity (3.32) and increasing access to digital technology (3.28) are the key strategies for reducing the challenges to the formation of digital education environments.

In addition, the study participants agreed to all the questionnaire items 21-30, indicating that the respondents perceive them as the key measures for reducing the problems confronting the development of digital education for the training of future special educators.

Table 4

Mean rating and standard deviation of study participants on measures reducing the problems confronting the formation of digital education

S/N	The measures reducing the problems confronting the formation of digital education include:	Mean	SD	Decision
1.	Provision of quality training on digital tools	3.40	0.60	Agreed
2.	Increasing access to digital technology	3.28	0.65	Agreed
3.	Provision of maximum technical support	3.45	0.58	Agreed
4.	Improving internet connectivity	3.32	0.62	Agreed
5.	Making digital education environments user-friendly	3.20	0.68	Agreed
6.	Provision of institutional support	3.35	0.64	Agreed
7.	Reducing the high cost of technology	3.18	0.70	Agreed
8.	Increasing collaboration among special educators	3.12	0.71	Agreed
9.	Provision of adequate incentives for special educators	3.05	0.73	Agreed
10.	Regular evaluation of digital education environments	3.22	0.67	Agreed
11.	Grand mean for research question three	3.26	0.66	Agreed

Research Question 4: In what ways can the prospects of a digital education environment be used for improving the training of special educators?

Data in Table 5 indicated that the future of the digital education environment would enhance special educators' skills, provide specialized learning experiences, provide opportunities for professional development, increase access to education for students with special educational needs, and enhance the quality of education for learners with special needs as they obtained mean scores of 3.40, 3.38, 3.30, 3.35 and 3.38,

respectively.

However, the participants of the study accepted all items presented to them, with a grand mean score of 3.27 and a standard deviation of 0.65, indicating that they perceive that the prospect of a digital education environment will improve the training quality of future special educators and also foster the academic achievement of learners with special education needs.

Table 5

Mean rating and standard deviation of study participants regarding how the prospects of a digital education environment can be used for improving the training of future special educators

S/N	The prospects of a digital education environment can be used for improving the training of special educators through:	Mean	SD	Decision
12.	Enhancing students' academic engagement	3.23	0.66	Agreed
13.	Providing specialized learning experiences	3.38	0.61	Agreed
14.	Increasing access to education for students with special needs	3.30	0.64	Agreed
15.	Improving communication between parents and educators	3.20	0.68	Agreed
16.	Providing opportunities for professional development	3.35	0.63	Agreed
17.	Enhancing the quality of education for learners with special needs	3.28	0.65	Agreed
18.	Increasing students' academic motivation	3.18	0.70	Agreed
19.	Providing opportunities for collaboration among learners	3.12	0.71	Agreed
20.	Improving the academic achievement of learners	3.22	0.67	Agreed
21.	Enhancing special educators' skills	3.40	0.60	Agreed
22.	Grand mean for research question four	3.27	0.65	Agreed

Discussion

The findings clearly reveal that respondents perceived that digital education environments foster the training of future special educators. The findings imply that study participants agreed on the relevance of digital education environments in providing maximum opportunities for professional development, ensuring that training resources are accessible and also enhancing flexible learning experiences. The grand mean score indicates that respondents perceive that digital educational environments can support the quality of training for future special educators.

To support this result, Ainscow (2020) and Fernandez-Batanero et al. (2022), agreed that when digital education is perceived as effective, it promotes effective teaching and learning and also encourages an all-inclusive education system. In another related finding, Falloon (2020) maintained that integrating digital technologies, content knowledge, and pedagogy is a significant way to promote effective teaching and learning.

This finding is also consistent with those of Claro et al. (2024), Zhang et al. (2025) and Kerimbayev et al. (2023) which stress that educators acquire the relevant knowledge and skills needed to promote quality pedagogical practices through well-designed digital education environment. Interpreting these findings through the lens of the TPACK framework, digital learning environments could be used to enhance the integration of pedagogical, content, and technological knowledge, as they would aid future special educators in effectively developing the knowledge, skills, and competencies required for teaching effectiveness.

Secondly, the results vividly reveal the perceived problems affecting the formation of digital education environments, especially for future special educators, such as technical issues, poor access to digital technology, and inadequate training in digital tools. These findings are consistent with those of Guillén-Gámez and Mayorga-Fernández (2020) and Dias-Trindade et al. (2020), that competence, autonomy, and relatedness enhance academic engagement. In a similar vein, Yakavets et al. (2023) supported that limited access to digital technologies and inadequate training hinder the effective integration of digital tools in education in Kazakhstan context. Moreover Kerimbayev et al. (2023) identified a significant digital divide between national and regional universities in Kazakhstan.

The findings also identified key strategies that could be perceived to mitigate problems impeding the formation of digital education environments. In this regard, the findings reveal that increasing access to digital

technology, training future special educators on digital tools, and providing technical support are perceived as the main strategies for reducing impediments.

It is necessary to integrate emerging technologies with content knowledge and pedagogical practices to promote quality teaching and learning with evidence from pedagogical universities in Kazakhstan as noted by Dang et al. (2024); De Obesso et al. (2023) and Gallego Joya et al. (2025). In addition, educators are more likely to use digital educational environments effectively when adequate technical support and high-quality training are provided.

Interestingly, the findings by Vuorikari et al. (2022) supported the urgent need to provide educators with training and support to enable them to integrate digital technologies effectively across the entire education system. These strategies could foster the autonomy, relatedness and competence of future special educators through the lens of Self-Determination Theory. Through training and technical support, educators could enhance their competence in using digital tools and resources and support their autonomy in pedagogical practices. Finally, the findings show that the prospect of digital education environments could be perceived to enhance future special educators' training. This result is in line with the findings by Ferrer et al. (2022) that academic engagement and motivation are more likely to improve through relatedness, competence and autonomy. In another view, Bravo-Agapito et al. (2021) noted that educators acquire the knowledge and skills to use digital resources and tools effectively through well-designed digital education environments.

The findings by Cabero-Almenera (2022) agreed that future special educators perceive digital environments as capable of providing specialized learning experiences, including assistive technologies for students with special education needs. Thus, it is worth noting that respondents perceived that prospect of digital education environments could improve training of future special educators. It is important to reiterate that these are perceptions; as this study is descriptive and survey-based using only mean and standard deviation, it does not prove effectiveness on student performance. Causal inferences cannot be made. The results of this research have important implications for educational stakeholders, such as students, educators, and educational policymakers. The findings of this study indicated that integrating digital education environments into future special educators' training promotes the quality of teaching and learning and improves students' academic performance.

It therefore becomes necessary for educational stakeholders to prioritize the formation of a well-designed digital education environment that encourages inclusivity, flexibility, accessibility in academic activities and also creates many opportunities for professional development. In addition, reducing the problems associated with the formation of digital education environments, including inadequate training and poor access to digital technology, is necessary to ensure that the integration of digital technology in the teaching-learning process is a reality.

The findings also highlighted that providing maximum training on how to use modern technologies, increasing access to digital technologies, and providing technical support are the most effective measures for reducing impediments to the formation of digital education environments. Imperatively, the successful implementation of digital education environments will go a long way toward transforming the education system and improving the lives and educational outcomes of learners with special educational needs.

Conclusion

This study investigated perceptions regarding formation of digital education environments for future special educators' training in the Republic of Kazakhstan with emphasis on key challenges and prospects. The findings show that respondents perceived that digital education environments can support high-quality training for future special educators, provide opportunities for professional development, and ensure that training resources are easily accessible. The results also indicated perceived key factors impeding the formation, such as the absence of training on technical issues and digital tools, as well as poor access to digital technologies. It was perceived that integrating digital education environments into traditional teaching approaches could create a personalized, effective academic environment.

The findings of this research have practical implications for higher education institutions, curriculum developers, and policymakers involved in special educator training. Higher education institutions could use these perceived challenges to develop programs geared towards promoting digital competency among future special educators. Curriculum planners could be motivated to effectively integrate digital environments into

the curricula of special education, while educational policymakers could allocate resources to improve access to digital technologies and address technical issues. However, the limitations of the study may affect the generalizability of the findings.

There is no study without limitations. The study was a descriptive survey measuring perceptions only, not actual competence or student outcomes; hence, causal claims cannot be made. The study population of 239 future special educators from three pedagogical universities in Kazakhstan, affects generalizability beyond similar contexts. The study relied on self-reported questionnaire; use of interviews, checklists, and observation could have improved depth. Moreover, no inferential statistics were applied; only mean and standard deviation were used. Future studies could adopt mixed methods, larger samples across Central Asia, and longitudinal designs to examine long-term impact of digital environments on actual digital competencies and outcomes of special education programs.

Conflict of Interest Statement

The authors declare no potential conflicts of interest regarding the research, authorship, or publication of this article.

Author Contributions

Kamilya Arinova: Conceptualization, Methodology, Collection and Curation, Supervision, Proofreading, Writing – Review and Editing. Yerkin Nurmakhanov: Literature Review, Data Analysis, Visualization, Writing – Original Draft, Final Editing.

Declaration of Use of Generative AI and AI-Assisted Technologies

The authors declare that no artificial intelligence (AI) tools were used in the preparation of this manuscript.

References

- Ainscow, M. (2020). Promoting inclusion and equity in education: lessons from international experiences. *Nordic Journal of Studies in Educational Policy*, 6(1), 7–16. <https://doi.org/10.1080/20020317.2020.1729587>
- Basilotta-Gómez-Pablos, V., Matarranz, M., Casado-Aranda, L.A. & Otto, A. (2022). Teachers' digital competencies in higher education: a systematic literature review. *Int J Educ Technol High Educ* 19, 8. <https://doi.org/10.1186/s41239-021-00312-8>
- Bond, M., Buntins, K., Bedenlier, S., Zawacki-Richter, O., & Kerres, M. (2020). Mapping research in student engagement and educational technology in higher education: A systematic evidence map. *International Journal of Educational Technology in Higher Education*, 17(1), 1–30. <https://doi.org/10.1186/s41239-019-0176-8>
- Bravo-Agapito, J., Romero, S. J., & Pamplona, S. (2021). Early prediction of undergraduate students' academic performance in completely online learning: A five-year study. *Computers in Human Behavior*, 115, 106595. <https://doi.org/10.1016/j.chb.2020.106595>
- Cabero-Almenara, J., Guillén-Gámez, F. D., Ruiz-Palmero, J., & Palacios-Rodríguez, A. (2022). Teachers' digital competence to assist students with functional diversity: Identification of factors through logistic regression methods. *British Journal of Educational Technology*, 53, 41–57. <https://doi.org/10.1111/bjet.13151>
- Cabero-Almenara, J., Gutiérrez-Castillo, J.-J., Palacios-Rodríguez, A., & Barroso-Osuna, J. (2020). Development of the Teacher Digital Competence Validation of DigCompEdu Check-In Questionnaire in the University Context of Andalusia (Spain). *Sustainability*, 12(15), 6094. <https://doi.org/10.3390/su12156094>
- Claro, M., Castro-Grau, C., Ochoa, J. M., Hinostroza, J. E., & Cabello, P. (2024). Systematic review of quantitative research on digital competences of in-service school teachers. *Computers & Education*, 215, 105030. <https://doi.org/10.1016/j.compedu.2024.105030>
- Dang, T., Phan, T., Vu, T., La, T., & Pham, V. (2024). Digital competence of lecturers and its impact on student learning value in higher education. *Heliyon*, 10, e37318. <https://doi.org/10.1016/j.heliyon.2024.e37318>

- De Obesso, M., Núñez-Canal, M., & Pérez-Rivero, C. A. (2023). How do students perceive educators' digital competence in higher education? *Technological Forecasting and Social Change*, 188, 122284. <https://doi.org/10.1016/j.techfore.2022.122284>
- Dias-Trindade, S., Moreira, J. A., & Ferreira, A. G. (2020). Assessment of university teachers on their digital competences. *QWERTY - Open and Interdisciplinary Journal of Technology, Culture and Education*, 15(1), 50–69. <https://doi.org/10.30557/QW000025>
- Falloon, G. (2020). From digital literacy to digital competence: The teacher digital competency (TDC) framework. *Education Tech Research Dev* 68, 2449–2472. <https://doi.org/10.1007/s11423-020-09767-4>
- Fernández-Batanero, J. M., Montenegro-Rueda, M., Fernández-Cerero, J., & García-Martínez, I. (2022b). Digital competences for teacher professional development. Systematic review. *European Journal of Teacher Education*, 45(4), 513–531. <https://doi.org/10.1080/02619768.2020.1827389>
- Fernández-Batanero, J. M., Montenegro-Rueda, M., Fernández-Cerero, J., & García-Martínez, I. (2022a). Assistive technology for the inclusion of students with disabilities: A systematic review. *Educational Technology Research and Development*, 70, 1911–1930. <https://doi.org/10.1007/s11423-022-10127-7>
- Fernández-Batanero, J.-M., Román-Graván, P., Reyes-Rebollo, M.-M., & Montenegro-Rueda, M. (2021). Impact of Educational Technology on Teacher Stress and Anxiety: A Literature Review. *International Journal of Environmental Research and Public Health*, 18(2), 548. <https://doi.org/10.3390/ijerph18020548>
- Ferrer, J., Ringer, A., Saville, K., Parris, M. A., & Kashi, K. (2022). Students' motivation and engagement in higher education: The importance of attitude to online learning. *Higher Education*, 83, 317–338. <https://doi.org/10.1007/s10734-020-00657-5>
- Gallego Joya, L., Merchán Merchán, M. A., & López Barrera, E. A. (2025). Development and strengthening of teachers' digital competence: Systematic review. *Contemporary Educational Technology*, 17(1), Article ep555. <https://doi.org/10.30935/cedtech/15744>
- Guillén-Gámez, F. D., & Mayorga-Fernández, M. J. (2020a). Identification of variables that predict teachers' attitudes toward ICT in higher education. *British Journal of Educational Technology*, 51, 1247–1262. <https://doi.org/10.1111/bjet.12933>
- Guillén-Gámez, F. D., & Mayorga-Fernández, M. J. (2020b). Quantitative-comparative research on digital competence in students, graduates and professors of Faculty Education: An analysis with ANOVA. *Education and Information Technologies*, 25(5), 4157–4174. <https://doi.org/10.1007/s10639-020-10160-0>
- Kerimbayev, N., Umirzakova, Z., Shadiev, R. & Jotsov, V. (2023) A student-centered approach using modern technologies in distance learning: a systematic review of the literature. *Smart Learn. Environ.* 10, 61. <https://doi.org/10.1186/s40561-023-00280-8>
- OECD (2021), *The State of School Education: One Year into the COVID Pandemic*, OECD Publishing, Paris, <https://doi.org/10.1787/201dde84-en>
- Taber, K.S. (2018). The Use of Cronbach's alpha when developing and reporting research instruments in science education. *Res Sci Educ* 48 (6), 1273–1296. <https://doi.org/10.1007/s11165-016-9602-2>
- Vuorikari, R., Kluzer, S. and Punie, Y. (2022) *DigComp 2.2, The Digital Competence framework for citizens- With new examples of knowledge, skills and attitudes (EUR31006EN)*. Publications Office of the European Union <https://data.europa.eu/doi/10.2760/115376>
- Yakavets, N., Winter, L., Malone, K., Zhontayeva, Z., & Khamidulina, Z. (2023). Educational reform and teachers' agency in reconstructing pedagogical practices in Kazakhstan. *Journal of Educational Change*, 24(4), 727–757. <https://doi.org/10.1007/s10833-022-09463-5>
- Yessenbekova, U., Aldabergenova, Z., Mamankul, A., Rakhymbayev, N., & Kutym, B. (2025). The impact of the digital format on Kazakh youth's information consumption and reading culture. *Publishing Research Quarterly*, 41, 113–131. <https://doi.org/10.1007/s12109-024-10012-6>
- Zhang, L., Yang, C., & Zheng, Y. (2025). Digital competence for sustainable education of pre-service teachers: A systematic literature review (2014–2024). *Frontiers in Psychology*, 16, 1710983. <https://doi.org/10.3389/fpsyg.2025.1710983>
- Zhao, Y., Llorente, A. M. P., & Gómez, M. C. S. (2021). Digital competence in higher education research: A

systematic literature review. Computers & Education, 168, 104212.
<https://doi.org/10.1016/j.compedu.2021.104212>

Information about authors

Kamilya Arinova – Master of Pedagogical Sciences, Doctoral Student of the Educational Programme “Special Pedagogy”, Abai Kazakh National Pedagogical University, Almaty, Republic of Kazakhstan; e-mail: arinova_k@buketov.edu.kz; ORCID iD: 0009-0003-9669-9345 (*corresponding author*).

Yerkin Nurmakhanov – Master of Technical Sciences, Lecturer, Department of Applied Mathematics and Computer Science, Karaganda National Research University named after Academician E. A. Buketov, Karaganda, Republic of Kazakhstan; e-mail: ereke.nurmahanov@gmail.com; ORCID iD: 0000-0002-2695-8394.

Received: 05.03.2026

Revised: 31.07.2026

Accepted: 30.09.2026

Published: 30.09.2026

Type of Article: Research Article

¹Lazura Kazykhankyzy, ¹Meruyert Seitova, ¹Ziyoda Khalmatova, ^{1,2}H. Eray Çelik*

¹ Foreign Languages Teaching Department, Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkistan, Republic of Kazakhstan

² Van Yüüncü Yıl University, Statistical Information Science, Van, Türkiye

EXPLORING EFL TEACHERS' ATTITUDES: ADAPTATION AND VALIDATION OF THE PROFESSIONAL IDENTITY SCALE IN KAZAKHSTAN

Abstract. The purpose of the present study was to adapt and validate the EFL Teacher Professional Identity Scale (EFLTPI) in the Kazakhstani context. The EFLTPI was completed by 470 English language teachers from diverse schools in the Turkistan region, Kazakhstan. A comprehensive approach was employed, utilizing a survey methodology to collect data using a Google Form questionnaire. Confirmatory factor analysis (CFA) was used to validate the measurement model, revealing satisfactory fit indices (Chi-square/sd = 1.862, RMSEA = 0.031, SRMR = 0.054, NFI = 0.933, NNFI = 0.968, CFI = 0.973, AGFI = 0.872). The scale demonstrated high reliability and validity, with Cronbach's alpha values varied between 0.86 and 0.94 for all factors. A second-level CFA indicated significant relationships among the factors, suggesting a unidimensional structure representing EFL teacher professional identity. Key findings highlighted high levels of perceived professional identity among teachers, particularly in serving as effective role models, creating effective teaching environments, possessing management skills, and imparting knowledge and experience. This study confirms the scale's applicability in Kazakhstan, providing a robust tool for assessing EFL teachers' professional identity and instilling confidence in the thoroughness of the research process.

Keywords: EFL teachers; professional identity; confirmatory factor analysis; Kazakhstan; survey method.

Introduction

Developing a professional identity is crucial for educators, particularly those engaged in teaching English as a foreign language (EFL), to progress and achieve greater success. A person's professional self-concept and function within the educational setting are defined by their ideas, attitudes, and practices, which collectively comprise their professional identity. EFL teachers' experiences, cultural settings, and work situations all influence their sense of self. These elements shape how educators perceive their responsibilities, interact with students, and develop as professionals. A professional identity is not static but changes over time due to ongoing reflection and modification. Given the frequent necessity to operate in multicultural and diverse settings, it is of the utmost importance for those teaching English as a Foreign Language (EFL) to develop a comprehensive understanding of their professional identities. The internalization of the standards and principles of the teaching profession, the adaptation to new educational regulations, and addressing the shifting needs of students are all components of this ongoing process. EFL teachers' professional identity enables them to maintain a coherent sense of self while navigating diverse institutional and cultural contexts, ensuring consistency in their interactions with students and instructional methods.

Further research in the Kazakhstani context is required to address the increased professional identity of EFL teachers. The reason for this is the impact of the factors above on instructional strategies, teacher motivation, and student results (Duff & Uchida, 1997; Norton, 2013). However, most current research is focused on Western and East Asian contexts, which has resulted in a significant knowledge gap regarding the construction of EFL professional identities in other countries, particularly in Central Asia. Investigating how Kazakhstani EFL teachers construct their professional identities may yield valuable insights into understanding the distinctive challenges and opportunities they encounter, thus contributing to the global discourse on teacher identity.

Perception and evaluation of EFL teachers' professional identities are crucial to improve instruction quality and student outcomes. The EFL Teacher Professional Identity Scale (EFLTPI) is a comprehensive instrument for evaluating the various dimensions of professional identity. The scale permits a sophisticated

comprehension of the various elements that contribute to a teacher's sense of professional identity, including their beliefs about teaching, their attitudes toward their students, and their practices in the classroom. Also, it enables researchers to comprehend better the intricate interrelationship between these factors and their collective influence on forming a teacher's professional identity.

This study aims to validate the EFLTPI in Kazakhstan by applying structural equation modeling (SEM) and confirmatory factor analysis (CFA). This will acquire distinctive insights into interrelationships between the constituent elements of professional identity and their influence on the efficacy of pedagogical practice in the specific geographical context. The study underscores the importance of assisting EFL teachers in professional development by examining the existing body of research on professional identity and its measurement. The validation of the EFLTPI in the Kazakhstani context is of particular importance, as it can assist in identifying the distinctive factors that contribute to the professional identity of EFL teachers in this region. Such factors may include the influence of local cultural and educational norms, the impact of national language policies, and the challenges and opportunities EFL teachers face in Kazakhstan. A deeper comprehension of these factors can furnish invaluable insights for policymakers, educational leaders, and teacher educators, thereby facilitating the formation of context-specific strategies to bolster the professional advancement of EFL teachers in Kazakhstan.

Literature Review

An EFL teacher's self-image and perspective of their position, duties, and tasks within the broader educational context is termed their "EFL professional identity." Many factors, including social influences, career experiences, educational background, and personal values, shape this identity. It is a complex notion that reflects the inherent challenges of teaching English to non-native speakers of the language.

As posited by Beijaard et al. (2004) identity is a dynamic and continuous process of reading and reinterpreting events. This perspective emphasizes that the professional identities of EFL teachers are malleable and evolve as they gain experience and encounter new challenges. Given their background and experiences as learners, teachers' views on teaching and learning significantly influence their engagement with students in the classroom (Richards, 2008).

Professional identity development is contingent upon participation in professional communities, wherein educators exchange practices, attitudes, and objectives (Wenger, 1998). Engagement in professional development activities, networking with colleagues, and affiliation with teaching organizations facilitate a sense of belonging and career advancement among EFL teachers. Furthermore, language use in social, cultural, and political contexts also shapes identity (Norton, 2013).

EFL teachers operate within many cultural contexts, and these factors, in conjunction with local perspectives on English, educational methodologies, and the language's position in society, collectively inform their professional identities. A further crucial element of an EFL professional identity is agency, which may be defined as the capacity of teachers to act with intention and make decisions within their professional environment (Lasky, 2005). EFL teachers exhibit agency by modifying their educational strategies, outlining their professional needs, and pursuing changes for professional development. It can be argued that maintaining a good professional identity is contingent upon having a sense of agency, especially in demanding educational settings. The interaction of personal, professional, and environmental elements highlights the intricacy of the EFL professional identity. The dynamic relationship between an individual and their surroundings shapes a teacher's identity, which always changes (Day et al., 2006).

The EFLTPI, developed by , assesses various dimensions of EFL teachers' professional identity. The scale includes 41 items distributed across 13 factors, such as creating a relaxed learning atmosphere, imparting knowledge and experience, having management skills, etc. Confirmatory factor analysis (CFA) is widely used to validate measurement models, ensuring their reliability and structural validity

The high standardized factor loadings (>0.70) and significant Cronbach's alpha, composite reliability (CR), and average variance extracted (AVE) values further confirm the scale's validity and reliability (Fornell & Larcker, 1981; Hair et al., 2014; Nunnally, 1967).

Professional identity in teaching is multifaceted, encompassing various interrelated factors. The significant correlations between the latent variables in the EFLTPI measurement model suggest strong interrelationships among these factors

It is essential to grasp how the identity of EFL professionals can be gauged to gain insight into the multifaceted aspects of teaching English in various contexts. Several scales and models have been developed globally to assess the various dimensions of this professional identity. Among the most notable of these are the Teacher Professional Identity Scale (TPIS), the Professional Identity of EFL Teachers Questionnaire (PIEFLTQ), the Teacher Identity and Professional Development (TIPD) Model, the EFL Teacher Identity Scale (EFL-TIS), the Language Teacher Identity (LTI) Framework, and the teaching Perspectives Inventory (TPI).

The TPIS developed by Beijaard et al. (2004) focuses on three main dimensions: subject matter expertise, didactic competence, and pedagogical expertise. The PIEFLTQ, developed by (Cheung, 2008), comprises items that assess teachers' beliefs, self-perceptions, and attitudes toward their profession. The TIPD model, proposed by Gee (2000), offers a sociocultural perspective on teacher identity. The EFL-TIS, developed by Varghese et al. (2005), captures the specific aspects of EFL teachers' professional identity. The LTI framework proposed by Barkhuizen and Consoli (2021), offers a narrative approach to understanding teacher identity. The TPI, developed by Pratt et al. (2000), identifies the dominant teaching perspectives of teachers. Pennington and Richards (2016) employ a range of scales and qualitative methods to investigate the professional identity of EFL teachers in diverse contexts. Trent (2013) used narrative inquiry to investigate the professional identity of novice EFL teachers in Hong Kong. Xu (2014) employed the PIEFLTQ to study the professional identity of Chinese EFL teachers.

These studies highlight developing and utilizing various scales and models for measuring EFL professional identity. These scales have enabled researchers to comprehend better the multifaceted nature of EFL teachers' professional identity and its influence on pedagogical practices, job satisfaction, and career advancement.

The measurement of EFL professional identity is grounded in several theoretical frameworks that provide a comprehensive understanding of the complex nature of teacher identity. These frameworks encompass sociocultural theory, identity theory, reflective practice, and narrative inquiry.

Sociocultural theory, based on the work of , highlights the role of social interaction and cultural context in developing professional identity. As articulated, identity theory provides a comprehensive lens to examine how individuals perceive and construct their professional identities. The frameworks provide a foundation for developing and applying diverse scales, ensuring their capacity to capture the multifaceted aspects of professional identity. Forming an EFL professional identity is not solely contingent upon individual beliefs and experiences; it is also significantly shaped by cultural and contextual factors. Such factors include societal attitudes toward the English Language, educational policies, and institutional cultures. EFL teachers frequently report a heightened sense of professional status and importance in contexts where English is highly valued. Conversely, EFL teachers may encounter obstacles in asserting their professional identity in societies where English is viewed with ambivalence or resistance. Educational policies and curriculum standards represent pivotal contextual factors that influence the formation of EFL professional identity. Policies that emphasize the importance of English and allocate resources for EFL programs can reinforce teachers' professional identity by providing them with the requisite support and recognition.

The culture of educational institutions significantly impacts the professional identity of EFL teachers. Institutional cultures conducive to collaboration, professional development, and teacher autonomy can reinforce this identity. The broader sociopolitical environment also profoundly influences the formation of EFL professional identity. EFL teachers may encounter conflicting expectations and pressures in contexts where political ideologies influence language policies.

Cultural norms and values about teaching and learning also influence the formation of an EFL professional identity. EFL educators are likely to possess a robust sense of professional identity and pride in their work in cultures that hold teachers in high esteem. The immediate classroom context, including the diversity of students and the resources available, also influences EFL teachers' professional identity.

Li and Xu (2023) explored the growth of EFL (English as a Foreign Language) teachers' educational identity through learning, using Cultural-Historical Activity Theory (CHAT) as their theoretical framework. The article examines the two EFL Chinese teachers' learning processes and describes that contradictions and reflections are sources of teacher learning. Teachers form their identities through agency and negotiation with others. The authors place emphasis on the fluid and transformative character of the teacher identity and argue

that this situation contributes to the better understanding of teacher learning and identity development processes.

Existing research on EFL professional identity in Kazakhstan and similar contexts has provided valuable insights into the particular challenges and opportunities faced by EFL teachers in these regions. However, there is a shortage of studies in this area. Shamim (2008) conducted a comparative study on the professional identity of EFL teachers in Pakistan and Kazakhstan, elucidating the impact of socio-political factors and educational policies on teachers' perceptions of their professional roles. Sercu et al. (2005) conducted a comparative study involving EFL teachers from eight countries, including Kazakhstan, to understand how cultural and contextual factors influence professional identity. The results of these studies demonstrate that the professional identity of EFL teachers is a dynamic and evolving phenomenon shaped by educational reforms, language policies, and societal attitudes toward English. These studies emphasize the necessity of providing ongoing professional development, institutional support, and peer collaboration opportunities to enhance EFL teachers' professional identity and effectiveness.

Lunenburg (2017) emphasized the crucial role of teachers as role models in shaping professional identity. Similarly, Wyatt (2018) underscored the importance of self-efficacy and role modeling in enhancing teachers' professional identity and instructional practices. Jia and Derakhshan (2023) investigated the interrelationships of Chinese English teachers' professional identity, commitment to the teachings and the professional achievement. Their study of 300 teachers based on every region in the country shows that both professional identity and commitment are significant factors of professional success as well. The authors point out the effect of such factors as emotional attachment of teachers to their career on the classroom efficiency and argue that positive professional identity and commitment can support more successful teaching (Jia & Derakhshan, 2023).

Effective teaching environments are essential for fostering student engagement and facilitating language acquisition (Borg, 2003). EFL teachers' high perceptions of their ability to create supportive and interactive classrooms align with research emphasizing the importance of conducive learning atmospheres for effective language instruction. Classroom management is vital for maintaining order and optimizing instructional time, contributing to student achievement and classroom efficiency (Evertson & Weinstein, 2015; Marzano et al., 2003). The commitment to imparting knowledge and experience reflects teachers' dedication to linguistic and cultural education, which is crucial for comprehensive language teaching (Richards, 2008).

Respectful interactions are essential, but other dimensions may be more salient in teachers' daily practice (Day et al., 2006). Similarly, there is a need for balanced instructional approaches and cultural competence in EFL teaching (Ellis, 2009; Kramsch, 2014).

The scale's comprehensive framework captures the multifaceted nature of professional identity, offering valuable insights for research, professional development, and educational policy. By understanding and supporting the various dimensions of professional identity, stakeholders can contribute to EFL teachers' professional growth and effectiveness.

Research Method

In the present study a survey method was used to collect data from 470 English language teachers from a variety of schools in Turkistan. The survey was conducted using a Google form. The questionnaire was sent to the participants by e-mail, and they completed it voluntarily.

The "EFL teacher professional identity scale" was used in this study. Preliminarily, a scale reading study was conducted with 15 English teachers with different academic degrees to determine whether a scale with 41 items under 13 factors is a comprehensive tool for measuring the professional identity of English teachers in Kazakhstan. As a result of a reading study, it was decided that the factors and items included in the scale were appropriate for the Kazakhstani educational system, and an application study was conducted.

According to the original EFLTPI scale a total of 41 items were distributed under the 13 factors as follows: (1) creating a relaxed learning atmosphere 4 items; (2) having the tendency to impart knowledge and experience five items; (3) having respectful behavior three items; (4) having the ability to develop/select EFL materials four items; (5) having management skills four items; (6) having error correction skills three items; (7) having communication skills three items; (8) creating an effective teaching environment three items, (9) having the tendency to develop professionally two items; (10) familiarity with target language and culture two

items; (11) serving as an effective role model 3 items; (12) valuing L1 culture two items; and (13) being concerned about students' ability and development three items (Hashemi et al., 2021)

The data collected were analyzed using SEM analysis to examine the relationship between the components of the scale and their impact on teacher identity. SEM analysis was used to simultaneously estimate measurement and structural models, comprehensively analyzing the hypothesized relationships. The first step was evaluating the measurement model to ensure the reliability and validity of the instrument. Further, confirmatory factor analysis was performed to assess the fit of the measurement model to the data and to evaluate the convergent and discriminant validity of the construct. Finally, second-level CFA was carried out to determine the effects of the components on EFLTPI acceptance.

Demographic Information

This study involved 470 English language teachers from various schools in the Turkistan region, with different teaching experiences, genders, and ages. The following Table 1 presents the demographic information of the participants.

Table 1

Socio-demographic information

Variable	Category	Frequency	%
Gender	Male	63	13,5%
	Female	407	86,5%
Age	20-25	17	4%
	26-30	159	34%
	31-35	105	22%
	35 +	189	40%
Teaching experience	0-5 years	28	6%
	6-10 years	117	25%
	11-15 years	206	44%
	16 +	119	25%

As can be seen from the table, the gender distribution of the participants is 86.5% female and only 13,5% male. The age distribution was found to be 4% between 20-25 years old, 22% between 31-35 years old, 34% between 26-30 years old, and 40% of participants over 35 years. The distribution of participants according to their teaching experience was as follows: 6% with 0-5 years, 25% with 6-10 years, another 25% with more than 16 years, and 44% with 11-15 years of experience.

Findings

The measurement model aims to reveal how and to what extent observed variables explain latent variables. Confirmatory factor analysis (CFA) is commonly used to develop measurement models and provides significant convenience. This method creates latent variables (factors) based on variables observed through a previously created model (Byrne, 2017; Raykov & Marcoulides, 2006; Schumacker, R & Lomax, 1996).

The scale used in this study consists of 41 items grouped under 13 latent variables. To determine whether the scale is statistically valid and reliable, a first-level CFA was conducted using LISREL 9.3.3.14 software.

The results of the first-level CFA revealed that the measurement model is statistically acceptable. The model's fit criteria are given in Table 2.

Table 2*Fit indices of the measurement model*

Fit Index	Value	Fit
Chi-square /sd	1.862	Good Fit
RMSEA	0.031	Good Fit
SRMR	0.054	Acceptable
NFI	0.933	Acceptable
NNFI	0.968	Acceptable
CFI	0.973	Good Fit
AGFI	0.872	Acceptable

A review of the standardized factor loadings of all items in the measurement model revealed that the minimum loadings of all items were above 0.70, indicating each item's sufficient power. Moreover, the standardized loadings of all items were found to be statistically significant at the 0.05 level, with the corresponding t-values presented in Table 3.

Table 3*CFA Results (Standard loading, t-value, Cronbach's Alpha, CR, AVE)*

Factor/ Item	Standard loading	t- value	Cronbach's Alpha	CR	AVE
FACTOR			0.918	0.918	0.736
RA1	0.87	-			
RA2	0.88	24.97			
RA3	0.85	23.96			
RA4	0.84	23.22			
FACTOR			0.943	0.944	0.770
IK1	0.85	-			
IK2	0.88	25.12			
IK3	0.93	27.93			
IK4	0.90	26.44			
IK5	0.83	22.80			
FACTOR			0.915	0.915	0.783
RB1	0.91	-			
RB2	0.89	28.12			
RB3	0.85	25.63			
FACTOR			0.936	0.937	0.788
DM1	0.90	-			
DM2	0.89	28.58			
DM3	0.91	29.98			
DM4	0.85	26.19			
FACTOR			0.862	0.865	0.616
MS1	0.82	-			
MS2	0.82	19.22			
MS3	0.73	16.59			
MS4	0.76	17.36			
FACTOR			0.897	0.900	0.751
ES1	0.91	-			
ES2	0.91	27.36			
ES3	0.78	21.41			
FACTOR			0.916	0.919	0.792
CS1	0.91	-			
CS2	0.93	30.72			
CS3	0.82	24.48			
FACTOR			0.904	0.906	0.763
TE1	0.83	-			
TE2	0.91	23.79			
TE3	0.88	23.07			
FACTOR			0.881	0.889	0.801
DP1	0.96	-			
DP2	0.83	14.42			
FACTOR			0.949	0.948	0.900
FL1	0.96	-			
FL2	0.94	25.65			

Factor/ Item	Standard loading	t- value	Cronbach's Alpha	CR	AVE
FACTOR			0.936	0.939	0.793
RM1	0.83	-			
RM2	0.93	26.32			
RM3	0.92	26.12			
RM4	0.88	23.89			
FACTOR			0.925	0.928	0.865
VL1	0.90	-			
VL2	0.96	22.43			
FACTOR			0.908	0.911	0.774
SA1	0.92	-			
SA2	0.89	27.24			
SA3	0.83	24.28		0.918	0.736

The Cronbach's alpha (CA), composite reliability (CR), and average variance extracted (AVE) values calculated for the measurement model provided strong evidence for the validity and reliability of the model. To ensure scale reliability, CA and CR values must be greater than 0.70 (Nunnally, 1978). In this study, the overall Cronbach's alpha value was calculated to be greater than 0.70 for each factor, with an overall alpha value of 0.944, as illustrated in Table 3. Similarly, CR values for all factors were above 0.70, and all AVE values were above 0.50 (Fornell & Larcker, 1981; Hair et al., 2014)

The following Table 4 presents the correlations between the latent variables.

Table 4

CFA Results (Standard loading, t-value, Cronbach's Alpha, CR, AVE)

Factor / Item	Standard loading	t- value	Cronbach's Alpha	CR	AVE
FACTOR			0.918	0.918	0.736
RA1	0.87	-			
RA2	0.88	24.97			
RA3	0.85	23.96			
RA4	0.84	23.22			
FACTOR			0.943	0.944	0.770
IK1	0.85	-			
IK2	0.88	25.12			
IK3	0.93	27.93			
IK4	0.90	26.44			
IK5	0.83	22.80			
FACTOR			0.915	0.915	0.783
RB1	0.91	-			
RB2	0.89	28.12			
RB3	0.85	25.63			
FACTOR			0.936	0.937	0.788
DM1	0.90	-			
DM2	0.89	28.58			
DM3	0.91	29.98			
DM4	0.85	26.19			
FACTOR			0.862	0.865	0.616
MS1	0.82	-			
MS2	0.82	19.22			
MS3	0.73	16.59			
MS4	0.76	17.36			
FACTOR			0.897	0.900	0.751
ES1	0.91	-			
ES2	0.91	27.36			
ES3	0.78	21.41			
FACTOR			0.916	0.919	0.792
CS1	0.91	-			
CS2	0.93	30.72			
CS3	0.82	24.48			
FACTOR			0.904	0.906	0.763
TE1	0.83	-			
TE2	0.91	23.79			
TE3	0.88	23.07			
FACTOR			0.881	0.889	0.801
DP1	0.96	-			

Factor / Item	Standard loading	t- value	Cronbach's Alpha	CR	AVE
DP2	0.83	14.42			
FACTOR			0.949	0.948	0.900
FL1	0.96	-			
FL2	0.94	25.65			
FACTOR			0.936	0.939	0.793
RM1	0.83	-			
RM2	0.93	26.32			
RM3	0.92	26.12			
RM4	0.88	23.89			
FACTOR			0.925	0.928	0.865
VL1	0.90	-			
VL2	0.96	22.43			
FACTOR			0.908	0.911	0.774
SA1	0.92	-			
SA2	0.89	27.24			
SA3	0.83	24.28		0.918	0.736

All correlations between latent variables were found to be statistically significant. In summary, the results demonstrate that the subscales of the scale, which consist of questions to measure the EFL teacher's professional identity, are reliable and structurally valid.

Table 5*Correlation Matrix of Latent Variables*

	RA	IK	RB	DM	MS	ES	CS	TE	DP	FL	RM	SA	VL
RA	1												
IK	0.43	0.521											
RB	0.346	0.572	1										
DM	0.387	0.473	0.382	1									
MS	0.418	0.452	0.470	0.518	1								
ES	0.381	0.378	0.398	0.442	0.591	1							
CS	0.341	0.491	0.421	0.463	0.413	0.543	1						
TE	0.470	0.422	0.431	0.418	0.523	0.485	0.521	1					
DP	0.299	0.247	0.284	0.369	0.395	0.354	0.373	0.398	1				
FL	0.246	0.425	0.319	0.351	0.339	0.305	0.349	0.391	0.344	1			
RM	0.295	0.534	0.497	0.481	0.386	0.382	0.494	0.458	0.330	0.477	1		
SA	0.379	0.380	0.347	0.371	0.438	0.421	0.396	0.488	0.311	0.352	0.464	1	
VL	0.196	0.439	0.441	0.401	0.384	0.277	0.370	0.363	0.406	0.566	0.616	0.389	1

Once the first-level factor analysis established that each item within the scale's sub-dimensions accurately and consistently measures the corresponding dimensions, a second-level confirmatory factor analysis should be conducted to determine the effects of these sub-dimensions on the scale. In this section of the study, the second-level measurement model was constructed by measuring the effect of each sub-dimension on the scale.

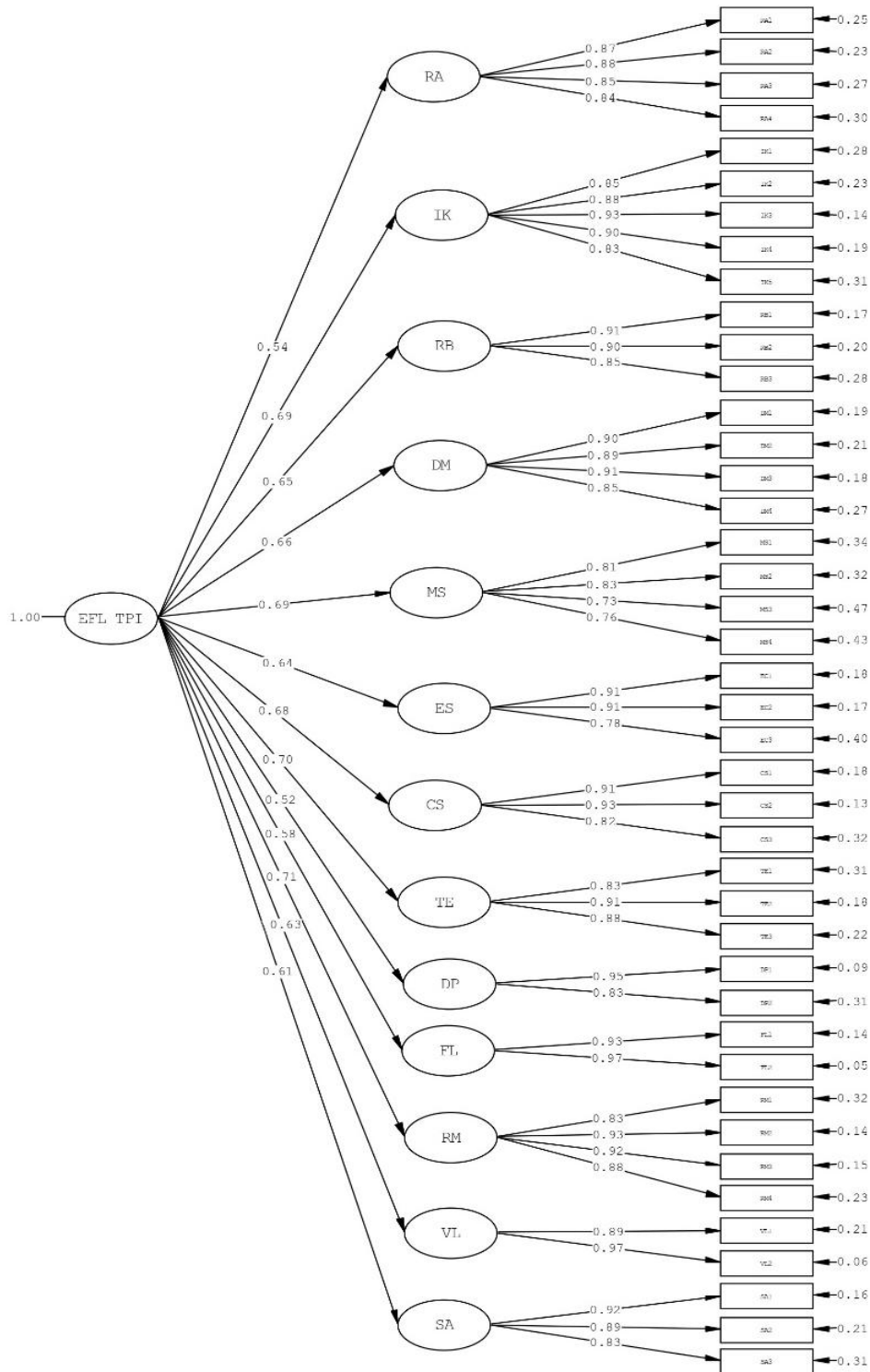
The results of the analysis provided statistical evidence that the model was the correct one. These results are presented in Table 6.

Table 6*Fit indices of the measurement model*

Fit Index	Value	Fit
Chi-square /sd	1.862	Good Fit
RMSEA	0.043	Good Fit
SRMR	0.054	Acceptable
NFI	0.916	Acceptable
NNFI	0.968	Acceptable
CFI	0.959	Good Fit
AGFI	0.854	Acceptable

This model was employed to examine the relationship between the EFLTPI latent variable and its sub-dimensions, which are suggested to constitute the scale in theory. The second-level measurement model is illustrated in Figure 1.

Figure 1
Measurement model



As shown in Figure 1, the standardized regression coefficients are above 0.50 and range between 0.52 and 0.71. The latent variable that exerts the most significant influence on EFLTPI is the RM (serving as an

effective role model) sub-dimension, with a coefficient of 0.71. In contrast, the variable that exerts the most minor influence is DP (tending to develop professionally), with a coefficient of 0.52.

The effects of the remaining sub-dimensions on EFLTPI are as follows: 0.70 TE (creating an effective teaching environment), 0.69 MS (having management skills), 0.69 IK (having the tendency to impart knowledge and experience), 0.66 DM (having the ability to develop/select EFL materials), 0.65 RB (having respectful behavior), 0.64 ES (having error correction skills), 0.63 VL (valuing L1 culture), 0.61 SA (being concerned about students' ability and development), 0.58 FL (familiarity with target language and culture) and 0.54 RA (creating a relaxed learning atmosphere). All regression coefficients in the second-level measurement model were statistically significant ($p < 0.05$).

After determining the items' discrimination power and obtaining Cronbach's alpha (CA), CR, and AVE values, it was concluded that the scale's validity and reliability were statistically at the desired level.

Given that the EFL teacher professional identity inventory used in this study has been used in previous studies (Hashemi et al., 2021), a confirmatory factor analysis was conducted to test its validity. The structural models were obtained through CFA. The optimal model was identified through the use of fit indices. In the EFLTPI scale, two models were considered: a first-level and a second-level multi-factor model. The suitability of these models was then evaluated.

When the correlations between the factors were examined, a significant relationship was determined, indicating that the factors in the scale could be represented by a single structure, namely the EFL teacher professional identity scale. As a result, the measurement tool used within this study's scope has been proven valid. Additionally, the initial theory that the dimensions of the measurement tool can be represented in a single structure has been confirmed.

In addition, the findings of this study revealed a high level of perceptions of English language teachers regarding their professional identity as teachers. In particular, their perceptions regarding RM (serving as an effective role model), TE (creating an effective teaching environment), MS (having management skills), and IK (tending to impart knowledge and experience) were identified as being at a high level. Other factors related to teacher identity were found to be at moderate levels.

Discussion

This study's first-level confirmatory factor analysis (CFA) aimed to validate the EFL Teacher Professional Identity Scale (EFLTPI) measurement model. The measurement model's fit indices indicate an acceptable fit, with Chi-square/df = 1.862, RMSEA = 0.031, SRMR = 0.054, NFI = 0.933, NNFI = 0.968, CFI = 0.973, and AGFI = 0.872. These values suggest that the model meets the criteria for good fit (Hu & Bentler, 1999), corroborating previous studies utilizing CFA to validate similar scales (Byrne, 2017; Raykov & Marcoulides, 2006).

The standardized factor loadings for all items in the measurement model exceeded 0.70, demonstrating sufficient power and statistical significance at 0.05 (Hair et al., 2014). The reliability and validity of the measurement model were further supported by high Cronbach's alpha (CA), composite reliability (CR), and average variance extracted (AVE) values, all exceeding the recommended thresholds (Fornell & Larcker, 1981; Nunnally, 1978). These results affirm that the EFLTPI is a reliable and valid instrument for assessing EFL teachers' professional identity, aligning with the findings of Hashemi et al. (2021).

The validation of the EFLTPI through CFA is consistent with previous research on teacher professional identity scales. For instance, Hashemi et al. (2021) reported similar fit indices and reliability measures, reinforcing the robustness of the EFLTPI across different contexts. Moreover, studies by Wang et al. (2016) and Beijaard et al. (2004) have highlighted the importance of using reliable measurement models to capture the multifaceted nature of professional identity in EFL teaching.

The correlations between the latent variables in the measurement model were statistically significant, indicating strong interrelationships among the factors. This finding suggests that the various dimensions of professional identity are interconnected and collectively contribute to the overall construct. Similar interrelationships have been observed in other studies, where self-efficacy, professional commitment, and interaction with students and colleagues were significantly correlated (Day et al., 2006; Tschannen-Moran & Hoy, 2001).

The second-level confirmatory factor analysis (CFA) was employed to determine the effects of the sub-

dimensions on the EFLTPI. The fit indices for the second-level model (Chi-square/df = 1.862, RMSEA = 0.043, SRMR = 0.054, NFI = 0.916, NNFI = 0.968, CFI = 0.959, AGFI = 0.854) indicate an acceptable fit, further validating the model's structure. The standardized regression coefficients, ranging from 0.52 to 0.71, demonstrate significant effects of the sub-dimensions on the overall construct, with RM (serving as an effective role model) exerting the most significant influence (0.71) and DP (tending to develop professionally) the least (0.52).

The significant impact of the RM sub-dimension on professional identity aligns with findings from other research. For instance, Lunenburg (2017) emphasized the critical role of teachers serving as role models in shaping professional identity. Similarly, Wyatt (2018) highlighted the importance of self-efficacy and role modelling in enhancing teachers' professional identity and instructional practices.

The moderate influence of the DP sub-dimension suggests that professional development, while essential, may not be as strongly perceived as other aspects of professional identity. This finding is consistent with the work of Vangrieken et al. (2017), who noted that while continuous professional development is crucial, its perceived impact can vary depending on the context and the specific developmental activities undertaken.

The second-level measurement model was validated by the statistical significance of all regression coefficients ($p < 0.05$), thereby reinforcing the model's reliability. The results above demonstrate the interdependence among the sub-dimensions and their collective impact on the overarching concept of professional identity. The sub-dimension and the overall construct demonstrate robust correlations that align with the theoretical framework of . The framework proposes that interrelated elements constitute professional identity.

EFL teachers demonstrate a profound dedication to moral conduct and professional conduct, as evidenced by their high regard for the role of an effective role model (RM). This is a pivotal aspect in the context of EFL, given that teachers frequently serve as linguistic and cultural role models for their pupils (Duff & Uchida, 1997). The findings of Lunenburg (2017) underscored the influence of role modeling on student motivation and learning outcomes, aligning with the high RM ratings.

EFL teachers demonstrate high confidence in their ability to construct dynamic and supportive learning environments, as evidenced by their high rating on the Teaching Environment (TE) scale. This aspect has great consequences for language acquisition. It supports by Borg (2003) study, which emphasizes the role of a supportive classroom atmosphere in encouraging student participation and learning. The high TE scores indicate that EFL teachers attach great importance in creating learning-friendly environments to deliver effective language instruction.

The observation that EFL teachers hold their management skills in high regard indicates their capacity to plan and direct lessons in the classroom. Effective management of the learning environment and utilizing instructional time are essential for attaining these goals (Marzano et al., 2003). These findings are consistent with those of other studies highlighting the value of management abilities in raising student success and improving classroom effectiveness (Evertson & Weinstein, 2015).

The Imparting Knowledge and Experience (IK) factor evidences EFL teachers' dedication to sharing knowledge and promoting student learning. The high IK ratings demonstrate the value educators place on transferring linguistic and cultural information, which is crucial for comprehensive language understanding (Richards, 2008). The findings are consistent with Hashemi et al. (2021), who also identified that EFL teachers had high expectations for the transfer of information.

The level of respectful behavior (RB) demonstrated by the participants was perceived to be high, although it was not as strongly felt as observed in the conduct of RM, TE, MS, and IK. This suggests that, despite an awareness of the value of courteous communication, other aspects of EFL teachers' professional identities may be more prominent in their day-to-day work. This result is consistent with the findings of Day et al. (2006), who highlighted that an individual's professional identity encompasses a range of attitudes and actions that are accorded varying degrees of importance.

EFL teachers may prefer alternative instructional techniques over error correction, as indicated by their moderate perception of error correction skills (ES). The importance of error correction in language acquisition is widely acknowledged; however, its significance varies according to the teaching styles employed and the classroom dynamics (Ellis, 2009). This research underscores the importance of adopting a balanced approach

to teaching that incorporates error correction without compromising other key aspects of the learning process.

The modest assessment of familiarity with the target language and culture (FL) indicates that, although EFL teachers place a high importance on cultural understanding, other characteristics of their professional identity may be of greater significance. This result is consistent with the findings of earlier studies, which suggest that cultural familiarity is an important factor in the context of EFL but may be superseded by managerial and pedagogical abilities (Kramsch, 2014).

The study's findings underscore the importance of targeted professional development initiatives that address the overlooked facets of professional identity. To assist EFL teachers in developing a more comprehensive professional identity, professional development programs that emphasize improving error correction skills (ES) and respectful behavior (RB) may prove beneficial. Furthermore, programs encouraging immersion in the target language and culture (FL) can enhance teachers' pedagogical approaches and cultural competency.

It is imperative that educational policies acknowledge and encourage the diverse professional identities of EFL teachers. The comprehensive professional identity of EFL teachers may be enhanced by policies that facilitate collaboration among educators, provide resources for professional growth, and cultivate a conducive learning atmosphere (Day et al., 2006). Policies that recognize and honor educators for their efforts to set a good example for students and provide productive learning environments may also serve to reinforce their dedication and motivation.

Conclusion

The present study employed first- and second-level confirmatory factor analyses (CFA) to validate Kazakhstan's EFL Teacher Professional Identity Scale (EFLTPI).

The study's findings illustrate that EFL teachers in Kazakhstan perceive themselves as highly professional, particularly in their capacity to serve as role models, design productive learning environments, demonstrate managerial abilities, and disseminate information and expertise. These factors significantly influenced teachers' professional identities, underscoring their pivotal role in the educational context.

The suggestion of targeted professional growth topics is based on moderate impressions of other aspects, including proficiency with the target language and culture, error repair abilities, and courteous demeanor. Enhancing these elements will facilitate the development of a more comprehensive professional identity among EFL teachers, enhancing their efficacy and self-assurance in the classroom.

To sum up, the study's results support the validity and reliability of the EFLTPI in the Kazakhstani context, providing a valuable tool for assessing and understanding the professional identities of EFL teachers. The findings substantiate the necessity for sustained professional development and governmental initiatives that acknowledge and promote the multifaceted nature of teacher identity. Consequently, those involved in education may enhance EFL teachers' effectiveness and professional development, ultimately improving language learning outcomes.

This study's validation of the EFL Teacher Professional Identity Scale confirms its efficacy in assessing EFL teachers' professional identities. The strong correlations between the components and the validation of the single-structure model demonstrate the intertwined nature of professional identity. The positive perceptions held by EFL teachers of themselves as managers, knowledge transmitters, competent role models, and designers of successful learning environments reinforce the positive aspects of their professional identities.

These findings emphasize the importance of providing EFL teachers with the support they require to develop their professional identities. They also have significant implications for educational policy and professional development. Future studies should address this study's shortcomings and investigate the use of mixed methods, cross-cultural comparisons, and the evolution of professional identity over time. Investigating and supporting EFL teachers' professional identities can enhance their efficacy and, subsequently, the quality of students' educational outcomes.

Funding Information

This study did not receive any specific grants from public, commercial or non-profit funding organizations.

Conflict of Interest Statement

The authors declare that there is no conflict of interest in this study.

Author Contributions

Lazura Kazykhankyzy, Meruyert Seitova, Ziyoda Khalmatova, and H. Eray Çelik contributed equally to the conceptualization, methodology, data collection, formal analysis, interpretation of the findings, preparation of the original manuscript, review and editing, supervision, and finalization of the study. All authors participated equally in all stages of the research and approved the final version of the manuscript.

Declaration of Use of Generative AI and AI-Assisted Technologies

The authors declare that no artificial intelligence (AI) tools were used in the preparation of this manuscript.

References

- Barkhuizen, G., & Consoli, S. (2021). Pushing the edge in narrative inquiry. In *System* (Vol. 102). <https://doi.org/10.1016/j.system.2021.102656>
- Beijaard, D., Meijer, P. C., & Verloop, N. (2004). Reconsidering research on teachers' professional identity. *Teaching and Teacher Education*, 20(2), 107–128. <https://doi.org/10.1016/j.tate.2003.07.001>
- Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching*, 36(2), 81–109. <https://doi.org/10.1017/S0261444803001903>
- Byrne, D. (2017). Data collection: What are life history interviews? *SAGE Research Method*.
- Cheung, H. Y. (2008). Measuring the professional identity of Hong Kong in-service teachers. *Journal of In-Service Education*, 34(3), 375–390. <https://doi.org/10.1080/13674580802003060>
- Day, C., Kington, A., Stobart, G., & Sammons, P. (2006). The personal and professional selves of teachers: Stable and unstable identities. *British Educational Research Journal*, 32(4), 601–616. <https://doi.org/10.1080/01411920600775316>
- Duff, P. A., & Uchida, Y. (1997). The Negotiation of Teachers' Sociocultural Identities and Practices in Postsecondary EFL Classrooms. *TESOL Quarterly*, 31(3), 451. <https://doi.org/10.2307/3587834>
- Ellis, R. (2009). Corrective Feedback and Teacher Development. *L2 Journal*, 1(1). <https://doi.org/10.5070/l2.v1i1.9054>
- Evertson, C. M., & Weinstein, C. S. (2015). Classroom Management as a Field of Inquiry. In *Handbook of Classroom Management*. <https://doi.org/10.4324/9780203874783.ch1>
- Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. In *Source: Journal of Marketing Research* (Vol. 18, Issue 1).
- Gee, J. P. (2000). Identity as an Analytic Lens for Research in Education. <http://re.aera.net>
- Hair, J., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate Data Analysis: Pearson New International Edition*. British Library Cataloguing-in-Publication Data.
- Hashemi, M. R., Karimi, M. N., & Mofidi, M. (2021). Developing and Validating an EFL Teacher Professional Identity Inventory: A Mixed Methods Study 1. In *MEXTESOL Journal* (Vol. 45, Issue 1).
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1). <https://doi.org/10.1080/10705519909540118>
- JIA, H., & DERAKHSHAN, A. (2023). Chinese English teachers' professional identity and commitment and their associations with their professional success. *Porta Linguarum Revista Interuniversitaria de Didáctica de Las Lenguas Extranjeras*. <https://doi.org/10.30827/portalin.vi2023c.29625>
- Kramsch, C. (2014). Teaching foreign languages in an era of globalization: Introduction. *Modern Language Journal*, 98(1), 296–311. <https://doi.org/10.1111/j.1540-4781.2014.12057.x>
- Lasky, S. (2005). A sociocultural approach to understanding teacher identity, agency and professional vulnerability in a context of secondary school reform. *Teaching and Teacher Education*, 21(8), 899–916. <https://doi.org/10.1016/j.tate.2005.06.003>
- LI, X., & XU, J. (2023). Understanding tertiary EFL teacher learning and identity development: a cultural-

- historical activity theory perspective. *Porta Linguarum Revista Interuniversitaria de Didáctica de Las Lenguas Extranjeras*. <https://doi.org/10.30827/portalin.vi2023c.29652>
- Lunenburg, F. C. (2017). *Organizational Structure and Design* Fred C. Lunenburg Sam Houston State University. *Organization Management*, 1(I).
- Marzano, R., Marzano, J., & Pickering, D. (2003). *Classroom management that works: Research-based strategies for every teacher*. In Online Submission.
- Norton, B. (2013). Bonny Norton - Identity and Language Learning_ Extending the Conversation-Multilingual Matters (2013). In *Identity and Language Learning Extending the Conversation: Vol. 2nd edition*.
- Nunnally. (1967). *Nunnally on Reliability. Psychometric Theory*, 2nd editio.
- Nunnally, J. C. (1978). *Psychometric Theory*. 2nd Edition. Diangong Jishu Xuebao/Transactions of China Electrotechnical Society, 2.
- Pennington, M. C., & Richards, J. C. (2016). Teacher Identity in Language Teaching: Integrating Personal, Contextual, and Professional Factors. *RELC Journal*, 47(1), 5–23. <https://doi.org/10.1177/0033688216631219>
- Pratt, D. D., Collins, J., & Jarvis-Selinger, S. (2000). *Development and Use of The Teaching Perspectives Inventory (TPI) Pedagogical Cultures in Medical Education View project Understanding Outliers: An investigation of orthopaedic educators' perspectives on teaching View project*. www.TeachingPerspectives.com
- Raykov, T., & Marcoulides, G. A. (2006). On multilevel model reliability estimation from the perspective of structural equation modeling. *Structural Equation Modeling*, 13(1). https://doi.org/10.1207/s15328007sem1301_7
- Richards, J. C. (2008). *Teaching Listening and Speaking From Theory to Practice*. www.cambridge.org
- Schumacker, R & Lomax, R. (1996). *A beginner's guide to structural equation modeling*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Sercu, L., Bandura, E., Castro, P., Davcheva, L., Laskaridou, C., Lundgren, U., del Carmen Méndez García, M., & Ryan, P. (2005). Foreign Language Teachers and Intercultural Competence: An International Investigation. In *Foreign Language Teachers and Intercultural Competence: An International Investigation*. <https://doi.org/10.1177/016146810610800820>
- Shamim, F. (2008). Trends, issues and challenges in English language education in Pakistan. *Asia Pacific Journal of Education*, 28(3), 235–249. <https://doi.org/10.1080/02188790802267324>
- Trent, J. (2013). From learner to teacher: practice, language, and identity in a teaching practicum. *Asia-Pacific Journal of Teacher Education*, 41(4), 426–440. <https://doi.org/10.1080/1359866X.2013.838621>
- Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: capturing an elusive construct. In *Teaching and Teacher Education* (Vol. 17).
- Tsui, A. B. M. (2007). Complexities of identity formation: A narrative inquiry of an EFL teacher. *TESOL Quarterly*, 41(4), 657–680. <https://doi.org/10.1002/j.1545-7249.2007.tb00098.x>
- Vangrieken, K., Meredith, C., Packer, T., & Kyndt, E. (2017). Teacher communities as a context for professional development: A systematic review. In *Teaching and Teacher Education* (Vol. 61, pp. 47–59). Elsevier Ltd. <https://doi.org/10.1016/j.tate.2016.10.001>
- Varghese, M., Morgan, B., Johnston, B., & Johnson, K. A. (2005). Theorizing Language Teacher Identity: Three Perspectives and Beyond. *Journal of Language, Identity & Education*, 4(1), 21–44. https://doi.org/10.1207/s15327701jlie0401_2
- Vygotsky, L. (1978). *Lev S. Vygotsky - Mind in Society_ The Development of Higher Psychological Processes*-Harvard University Press (1978).
- Wang, L. Y., Jen-Yi, L., Tan, L. S., Tan, I., Lim, X. F., & Wu, B. S. (2016). Unpacking high and low efficacy teachers' task analysis and competence assessment in teaching low-achieving students in secondary schools. *Australian Educational Researcher*, 43(2), 165–183. <https://doi.org/10.1007/s13384-015-0196-x>
- Wenger, E. (1998). *Communities of Practice: Learning, Meaning, and Identity*. *Systems Thinker*, 9. <https://doi.org/10.2277/0521663636>
- Wyatt, M. (2018). Language teachers' self-efficacy beliefs: A review of the literature (2005-2016). *Australian Journal of Teacher Education*, 43(4). <https://doi.org/10.14221/ajte.2018v43n4.6>

Xu, Y. (2014). Becoming researchers: A narrative study of Chinese university EFL teachers' research practice and their professional identity construction. *Language Teaching Research*, 18(2), 242–259. <https://doi.org/10.1177/1362168813505943>

Information about authors

Lazura Kazykhankyzy - PhD, Assistant Professor, Foreign Languages Teaching Department, Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkistan, Republic of Kazakhstan; e-mail: lazura.kazykhankyzy@ayu.edu.kz; ORCID iD: 0000-0002-4155-6430.

Meruyert Seitova - PhD, Associate Professor, Head of the Foreign Languages Teaching Department, Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkistan, Republic of Kazakhstan; e-mail: meruyert.seitova@ayu.edu.kz; ORCID iD: 0000-0002-0735-2469.

Ziyoda Khalmatova - PhD, Senior Lecturer, Foreign Languages Teaching Department, Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkistan, Republic of Kazakhstan; e-mail: ziyoda.khalmatova@ayu.edu.kz; ORCID iD: 0000-0001-7695-5282.

H. Eray Çelik - PhD, Professor, Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkistan, Republic of Kazakhstan, and Van Yüzüncü Yıl University, Van, Türkiye; e-mail: ecelik@yyu.edu.tr; ORCID iD: 0000-0001-7490-8124 (*corresponding author*).

Received: 22.03.2026

Revised: 11.06.2026

Accepted: 18.09.2026

Published: 30.06.2026

Type of Article: Research Article

¹Aziza Zhunusbekova, ²Samal Askarkyzy*, ³Bakhyt Alipova

¹ Abai Kazakh National Pedagogical University, Almaty, Republic of Kazakhstan

² Kazakh National University of Arts named after K. Baiseitova, Astana, Republic of Kazakhstan

³ University of Kentucky, Kentucky, USA

AI ALGORITHMS FOR OPTIMIZING UNIVERSITY RESOURCE MANAGEMENT

Abstract. This article examines the potential of AI algorithms for optimizing university resource management in the context of the digital transformation of higher education. Modern approaches to the intellectualization of management processes, based on the use of optimization algorithms and machine learning methods, are analyzed. Attention is paid to heuristic and meta-heuristic approaches such as genetic algorithms, particle swarm modeling, and ant colony modeling, which are used to allocate educational, scientific, and human resources at universities. The research methodology includes a combination of complementary scientific methods: analysis of scientific publications and regulations, a comparative analytical approach, algorithmic modeling and optimization methods, statistical data processing, and conceptual modeling. The empirical basis of the study is data from 11 universities located in the cities of Astana and Almaty. The results demonstrate the significant potential of using AI algorithms to improve the efficiency of university resource management. The use of intelligent analysis and optimization methods enables more informed management decisions, promotes the rational allocation of educational, research, and human resources, and improves the quality of educational activities in the context of the digitalization of higher education.

Keywords: artificial intelligence, optimization algorithms, university resource management, digital transformation of education, genetic algorithms, particle swarm optimization, ant colony optimization.

Introduction

The Republic of Kazakhstan is consistently advancing in the sphere of digital transformation and the integration of artificial intelligence (AI) technologies. The country is gradually transitioning from isolated pilot projects toward a comprehensive system of state regulation and the widespread adoption of AI-based solutions. This strategic direction was officially formalized through Resolution №592 of the Government of the Republic of Kazakhstan adopted in 2024, which approved the Concept for the Development of Artificial Intelligence for the period 2024–2029. The Concept is a strategic document aimed at forming a national AI ecosystem through the development of infrastructure, data, human capital, and legal regulation, which should ensure the implementation of technologies in the economy and public administration. A special emphasis is placed on education, which envisages the large-scale training of specialists, the integration of AI into curricula, and the creation of a research environment based on universities, including access to international educational platforms and the development of initiatives such as AI- Sana (Government of the Republic of Kazakhstan, 2024). A significant step in this direction was the adoption of the Law «On Artificial Intelligence», which entered into force on January 18, 2026, demonstrating the formation of a regulatory framework for regulating the development and use of AI technologies (Republic of Kazakhstan, 2025). Furthermore, the institutional development of the industry is ensured by the creation of a specialized government agency – the Ministry of Artificial Intelligence and Digital Development of the Republic of Kazakhstan – overseeing issues of digital development and artificial intelligence (Republic of Kazakhstan, 2025).

The declaration of 2026 as the Year of Digitalization and Artificial Intelligence underscores the strategic importance of this sector for the country's socioeconomic development (President of the Republic of Kazakhstan, 2026). At the same time, technological infrastructure is actively developing: computing power based on GPU clusters is being implemented, and a national AI platform is being formed, providing access to government data and creating the conditions for the development of the concept of «sovereign artificial intelligence». The creation of this platform is aimed at developing an integrated AI development ecosystem, digitalizing data processing processes, and improving public administration mechanisms. Its implementation

is aimed at improving the quality of public services and strengthening Kazakhstan's competitiveness in the global artificial intelligence space. The practical application of AI already covers such areas as healthcare, transportation, and security (National Information Technologies JSC).

Particular attention is being paid to training personnel and developing human capital. Artificial intelligence programs are being implemented in the education system, including the adaptation of international educational practices, and national standards for integrating AI into higher education are being developed. These initiatives are aimed at training specialists capable of effectively applying digital technologies in various sectors of the economy, including higher education. By the end of 2025, more than 900,000 citizens had completed training, acquiring basic and applied competencies for work and study in the context of the rapid implementation of AI (Ministry of Digital Development, Innovations and Aerospace Industry of the Republic of Kazakhstan, 2026). These changes have also affected school education: starting in the 2025–2026 academic year, elements of artificial intelligence have been introduced into the subjects «Digital Literacy» and «Computer Science». Online Day of AI courses have been developed for secondary school students, and specialized professional development courses are available for teachers (Ministry of Education of the Republic of Kazakhstan, 2025). The Alem.ai center, opened in 2024, is a national platform for the development of artificial intelligence, focused on education, research, and innovation, bringing together pupils, entrepreneurs, and scientists. The project envisages the annual training of up to 1,000 specialists, the launch of 100 start-ups, support for 10 research initiatives, and the involvement of up to 10,000 participants by 2030, which is aimed at creating a sustainable AI ecosystem in Kazakhstan (Astana Hub, 2026). Creation of the AI SilkNet center – a national platform for predictive analytics and technological foresight – is aimed at developing scientific research, forming a unified forecasting system and strengthening interaction between the state, science and industry, which reflects the transition to development management based on data and intelligent technologies (Ministry of Digital Development, Innovations and Aerospace Industry of the Republic of Kazakhstan, 2026).

According to the Government AI Readiness Index 2025 prepared by Oxford Insights, Kazakhstan ranked 60th out of 195 countries, reflecting the continued development of its national AI ecosystem. The country demonstrates a high level of readiness for the implementation of artificial intelligence in public administration, particularly in the area of digital government services, as evidenced by its results in the Public Sector Adoption. The achieved results are due to the development of digital infrastructure, the institutional framework, and strategic state policy in the field of AI. At the same time, Kazakhstan's further progress is associated with the need to strengthen the commercialization of technologies, develop the private sector, and expand the practical application of AI in the economy (Ministry of Digital Development, Innovations and Aerospace Industry of the Republic of Kazakhstan, 2025). The achieved level of development of digital infrastructure and institutional readiness for the implementation of artificial intelligence in the Republic of Kazakhstan creates the preconditions for its active use at the level of educational organizations. In this context, the use of AI algorithms for optimizing university resource management, ensuring increased efficiency in the use of material, human, and information resources, as well as ensuring the sustainable development of the educational organization in the context of digital transformation, is of particular importance.

Literature Review

Currently, a large number of organizations across various sectors are implementing or are interested in implementing artificial intelligence (AI) technologies (Füller et al., 2022). Industries such as telecommunications, engineering, financial services, customer service, healthcare, pharmaceuticals, and educational institutions are at the forefront of AI tool adoption (Perifanis & Kitsios, 2023). The original goal of AI was to develop computer programs with capabilities similar to human intelligence. AI methods recreate the thought process and its outcome in an even more modest form (Dai et al., 2020). The rapid pace of digitalization, and in particular the adoption of AI, are changing the face of organizations. AI has changed the way organizations operate and has become part of corporate strategy. Universities today operate in a data-driven environment, where decision-making is increasingly based on the analysis of large amounts of information. At the same time, the further evolution of digital management involves not only the accumulation and analysis of data but also the use of intelligent tools for its processing, ensuring a higher level of management efficiency. In this context, the need to rethink traditional models of educational management is growing. As Daniela notes, the concept of «smart pedagogy» is being shaped by the challenges of the big data

era and the intellectualization of the educational environment. Established management practices and classical approaches to organizing the educational process are gradually losing their relevance, as they do not fully meet the increasingly complex needs of students and the demands of modern society, characterized by highly dynamic change.

The next stage of this transformation is the transition to the intellectualization of management processes. In the era of 5.0, characterized by the intensive development of information and communication technologies, artificial intelligence is becoming a key tool for technological and socioeconomic transformation. In the field of educational organization management, artificial intelligence technologies create the preconditions for improving the validity of management decisions, optimizing organizational processes, and enhancing the quality of educational activities. Optimizing the use of artificial intelligence technologies in education management opens up significant opportunities for improving institutional efficiency. At the same time, the effective allocation of academic, scientific, and personnel resources remains one of the central challenges in higher education management. The quality of educational services, research productivity, and the competitive position of a university largely depend on how effectively these resources are managed. At the same time, the complex internal structure of universities together with rapidly changing external conditions significantly complicate the process of achieving optimal resource allocation, turning it into a multifaceted management task.

Under the conditions of ongoing digitalization, methods of optimization driven by artificial intelligence technologies are becoming increasingly relevant, which creates a demand for the continuous improvement and adaptation of approaches to their integration into educational environments. Such approaches involve the use of the professional potential of specialized university divisions, including academic writing centers, the emergence of new positions related to AI support within university libraries, as well as institutional governance strategies designed to encourage innovative practices in the implementation of AI technologies in teaching and learning processes.

The development of digital educational resources (study materials, teaching aids, and practical AI examples) that facilitate the updating of course content is crucial, as is supporting the professional development of faculty through training, internships, and flexible teaching load distribution. An additional element is the expansion of university libraries' resource base through the acquisition of specialized publications and databases, which together ensures more effective management of the educational institution's material, human, and information resources.

According to the Law of the Republic of Kazakhstan «On Artificial Intelligence» dated November 17, 2025 №230-VIII LRK, artificial intelligence is the ability of systems to reproduce the cognitive functions inherent in humans, providing results comparable to or superior to the results of their intellectual activity (Republic of Kazakhstan, 2025). According to Tambe et al., artificial intelligence (AI) is an important driver of technological progress, influencing the labor market, which requires understanding its impact on scientific research (Tambe et al., 2019). With the expanding use of artificial intelligence technologies, the potential for job losses arises, as a result of persistent social inequality, as well as improved quality of work and the emergence of new forms of employment. In this regard, artificial intelligence is becoming a key factor in the transformation of human resource management systems (Ng et al., 2025). At a practical level, these changes manifest themselves in the use of intelligent algorithms to optimize management processes and resource allocation in organizations, including educational institutions.

Intelligent optimization algorithms, including genetic algorithms (GAs), particle swarm optimization (PSO), and ant colony optimization (ACO), can be applied to solving resource allocation problems for various types of universities. These algorithms are capable of handling complex, nonlinear, and multi-criteria models, making them suitable tools for resource management under conditions of high uncertainty. One of the priority areas for using such algorithms is optimizing the use of educational resources. This includes assigning instructors to subjects, creating class schedules, and assigning classrooms. The use of intelligent methods makes it possible to increase the consistency of the appointment of teachers, taking into account their professional specialization, the requirements of educational programs and the educational demand of students, as well as minimize conflicts in the schedule and increase student satisfaction. In particular, genetic algorithms can be used to optimize the assignment of instructors to courses. This model generates a set of possible solutions, which are evaluated against specified criteria – qualification, teaching load, and demand for the

discipline. Then, using selection, crossover, and mutation mechanisms, an evolutionary search is performed, resulting in an optimal or near-optimal assignment.

An equally significant application of optimization algorithms is scientific resource management. This includes distributing funding for scientific projects, assigning research tasks to research teams, and efficiently utilizing scientific infrastructure. Intelligent methods enable projects to be ranked by strategic importance, expected scientific impact, and feasibility, ensuring the informed allocation of resources. Thus, a particle swarm algorithm can be used to optimize the allocation of research funding. In this approach, potential solutions are modeled as particles in a multidimensional search space, and their positions are adjusted based on individual and global best results. This allows the system to gradually converge toward the optimal allocation of funds, taking into account multiple criteria.

In addition, artificial intelligence algorithms can be used to manage the university's human resources. It includes the processes of personnel selection, forecasting staff turnover, planning professional development. The use of intelligent models makes it possible to identify the most promising candidates, predict personnel risks and form effective employee development programs, which contributes to increasing organizational stability and staff satisfaction. According to Xu's research, models based on artificial intelligence algorithms provide more accurate and objective assessment of educational processes compared to traditional analysis methods. The use of optimization algorithms based on artificial intelligence allows for the development of more scientifically based and objective models for assessing the quality of teaching, increasing the reliability and validity of the results (Xu, 2025).

Various machine learning methods can be used to develop such models, including artificial neural networks (ANN), support vector machines (SVM), and decision trees (DT). These algorithms are able to identify complex, non-linear relationships between a variety of factors affecting the quality of teaching, including student reviews, professional characteristics of the teacher, the content and complexity of the course, as well as the educational outcomes of students. One possible approach is to build a model based on a multilayer feedforward neural network. The input layer can include a set of indicators characterizing teaching quality: student grades, teacher qualifications, subject characteristics, student academic performance, and other relevant parameters. Hidden layers enable the identification of complex relationships and patterns between these variables, while the output layer generates an integrated indicator of teaching quality that can be used in management decision-making. At the same time, the introduction of artificial intelligence into education management poses a number of significant challenges. One key issue is ensuring the privacy of students' personal data: the use of AI technologies requires strict adherence to information security standards and regulated data processing procedures. Equally important are the ethical aspects of using artificial intelligence, including issues of fairness, transparency of algorithms, and accountability for decisions made using automated systems. These factors require the development of regulatory and organizational mechanisms to govern the implementation of AI in the educational environment.

As Mahajan, Gupta, and Misra note, effective management in the digital age requires developed analytical and strategic competencies. In the context of university management, this means developing a new type of manager capable of working with digital tools, predictive models, and algorithmic decision support systems (Mahajan et al., 2022). According to Khatri et al. (2024) and Ratten and Jones (2023), the transformation of management processes in organizations requires a revision of traditional models of leadership and decision-making. These provisions are fully applicable to higher education institutions, where management structures are faced with the need to implement intelligent data analysis systems (Khatri et al., 2024; Ratten & Jones, 2023). At the same time, Greenland et al. (2022) point to the need for institutional readiness for digital transformation. The integration of generative artificial intelligence technologies into management processes requires the development of normative, ethical, and organizational regulatory mechanisms (Greenland et al., 2022). Consequently, despite the active discussion of AI in the educational process, the issue of its systemic implementation at the university governance levels – the dean's office, department, and rector's office remains insufficiently studied. The systematization of the areas of application of AI in university management is presented in Tab. 1.

Table 1*Application of artificial intelligence technologies in university resource management*

Direction of management	Management tasks	Algorithms used	Description of application	Expected effect
Educational resource management	Assignment of teachers, scheduling, and classroom management	GA (genetic algorithms), PSO, ACO	Optimizing faculty allocation based on qualifications, workload, and demand; minimizing scheduling conflicts	Increased assignment consistency, reduced conflict, and increased student satisfaction
Scientific resource management	Funding allocation, project management, infrastructure use	PSO	Optimizing the allocation of funds based on scientific significance, expected results, and feasibility of projects	Increasing funding efficiency and increasing scientific productivity
Human Resources Management	Recruitment, turnover forecasting, competency development	AI models (ML)	Analysis of professional potential, forecast of personnel risks, formation of development programs	Increasing organizational resilience and reducing personnel risks
Teaching quality assessment	Analysis of educational results and quality of teaching	ANN, SVM, DT	Identifying nonlinear relationships between teaching characteristics and learning outcomes	Improving the accuracy and objectivity of education quality assessment
Making management decisions	Data analysis and forecasting	AI, Big Data analytics	Using intelligent systems to support big data-based decisions	Improving the validity and effectiveness of management

Materials and Methods

The methodological basis of the study is a systems and analytical approach to managing university resources in the context of the digital transformation of higher education. The university is viewed as a complex socioeconomic system in which the management of educational, scientific, and human resources requires the use of modern data analysis methods and intelligent optimization algorithms. The paper uses the concept of Data-driven Management, which involves making managerial decisions based on the analysis of large amounts of data and the use of artificial intelligence algorithms. Within this approach, optimization algorithms are viewed as tools for intelligently supporting management decisions when allocating educational, scientific, and human resources.

The empirical base of the study consisted of 11 universities from Astana and Almaty. Four universities were located in Astana (36.4% of the sample), including one national university and three private universities, while seven universities were located in Almaty (63.6%), including three national and four private universities. Overall, the sample comprised four national universities (36.4%) and seven private universities (63.6%).

This distribution allows us to take into account the institutional features of educational resource management in universities of various types and provides institutional variation within the selected sample. To accomplish the objectives of the study, an integrated set of interrelated scientific methods was applied.

A review of academic sources and regulatory frameworks was used to examine contemporary approaches to artificial intelligence in educational administration and to identify key trends in the digital modernization of higher education.

A comparative analytical approach was used to compare the management models of the participating universities, with particular attention to differences in the allocation of academic, research, and personnel resources.

Algorithmic modeling and optimization techniques were used to assess Differential Evolution (DE), Particle Swarm Optimization (PSO), and Ant Colony Optimization (ACO) as tools for timetable generation, balancing faculty workloads, and improving the efficiency of research funding allocation.

Statistical data analysis was used to evaluate the effectiveness of management decisions and to identify relationships between resource allocation and indicators of educational activity.

Conceptual modeling was used to develop a framework for applying artificial intelligence algorithms to support university management decisions.

The integrated use of these methods has made it possible to identify the potential of artificial intelligence algorithms in improving the efficiency of university resource management and identify promising areas for their implementation in the higher education management system.

Results and Discussion

An analysis of the management practices of the universities included in the study sample revealed specific aspects of the use of digital technologies and artificial intelligence algorithms in the management of educational, scientific, and human resources. The findings indicate that while most universities already use various digital management systems, the level of implementation of intelligent optimization algorithms remains uneven.

The study found that 9 out of 11 universities (81.8%) use information systems to manage the educational process, including electronic scheduling systems, teacher and student databases, and workload management systems. However, full-fledged intelligent optimization algorithms are used much less frequently – only 4 universities (36.4%) use elements of analytical or predictive models when making management decisions.

One of the most relevant areas of application of artificial intelligence algorithms is the optimization of the curriculum. Analysis showed that traditional scheduling methods often lead to classroom conflicts and teacher overload. Modeling the application of optimization algorithms revealed that the use of the particle swarm optimization (PSO) and ant colony optimization (ACO) algorithms can reduce the number of scheduling conflicts by an average of 28–35% and also increase the efficiency of classroom utilization by approximately 20–25%.

Another important area is the optimization of the distribution of teaching load. Modeling results showed that the use of a differential evolution (DE) algorithm allows for the simultaneous consideration of several criteria – teacher qualifications, specialization, teaching load, and demand for the course. This algorithm can reduce the imbalance in teaching loads among teachers by approximately 18–22%, which contributes to a more equitable distribution of teaching hours. Also, significant opportunities for the use of artificial intelligence algorithms have been identified in the field of scientific resource management. Analysis has shown that the distribution of research funding at universities is often based on expert assessment and administrative decisions. The use of optimization algorithms, such as PSO, allows for the creation of more objective funding allocation models that take into account a variety of parameters, including scientific productivity, the strategic significance of projects, and expected research outcomes. Based on the modeling results, the efficiency of research fund allocation can be improved by 15–20%, which contributes to the growth of scientific productivity of universities. In addition, artificial intelligence algorithms can be used to assessing the quality of teaching and predicting educational outcomes. The use of machine learning models such as artificial neural networks, support vector machines, and decision trees allows for the identification of complex relationships between various educational process indicators. In particular, modeling has shown that the use of neural networks to analyze educational data can improve the accuracy of predicting student academic performance by up to 82–87%, while traditional statistical methods provide accuracy at the level 65–70%.

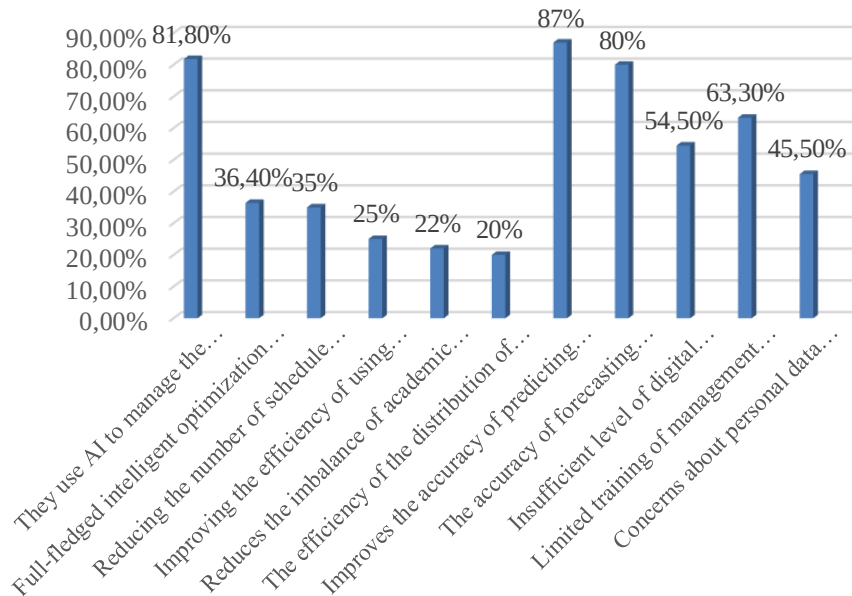
Special attention in the course of the study was paid to the management of human resources of universities. Artificial intelligence algorithms can be used to predict staff turnover, analyze the professional potential of faculty, and plan professional development. The use of intelligent models can reduce the risks of staff instability and improve the efficiency of human resource management. Based on the results of analytical modeling, the accuracy of personnel risk forecasting can reach 75–80%.

At the same time, the study results showed that the implementation of artificial intelligence algorithms in university management systems is accompanied by a number of organizational and institutional limitations. Among the main challenges, respondents noted an insufficient level of digital infrastructure (54.5% of universities), limited training of management personnel in the use of analytical tools (63.6%), and issues related to the protection of personal data (45.5%). Figure 1 clearly shows the relationship between the use of digital systems and artificial intelligence algorithms.

Accordingly, the results demonstrate the significant potential of artificial intelligence algorithms for optimizing university resource management. The use of intelligent analysis and optimization methods allows for more robust management decisions, improved allocation of educational, research, and human resources, and enhanced quality of university educational activities in the context of the digital transformation of higher education.

Figure 1

Use of digital systems and artificial intelligence algorithms in universities



A comparative analytical method was used to compare the management practices of the universities included in the study sample and identify differences in approaches to the allocation of educational, scientific and human resources. As Table 2 shows, national universities demonstrate a higher level of implementation of intelligent algorithms and analytical tools, while private universities predominantly employ traditional and partially digitalized approaches to resource management.

Table 2

Comparative analysis of management practices of national and private universities

Indicator	National universities (n=4)	Private universities (n=7)	Analytical interpretation
Implementation of digital management systems	100%	71.4%	National universities exhibit a more advanced level of institutional digitalization
Application of AI-based algorithms	50%	28.6%	The integration of AI technologies remains comparatively limited in private universities
Automation of timetable generation	High level (including partial AI integration)	Moderate level (primarily standard digital platforms)	Private institutions continue to rely more heavily on conventional management practices
Decrease in timetable conflicts	30–35%	25–30%	National universities demonstrate a stronger optimization effect
Efficiency of classroom utilization	80–85%	70–75%	Infrastructure resources are managed more effectively in national universities
Distribution of teaching workload	Optimized through DE algorithms	Partially coordinated	Private universities face a greater probability of uneven workload allocation
Model of research funding allocation	Algorithm-driven (PSO, analytical systems)	Administrative-expert approach	National universities employ more data-oriented decision-making mechanisms

Growth in efficiency of research resource management	18–20% increase	12–15% increase	The impact of optimization technologies is more evident in national universities
Use of analytics and predictive technologies	Broadly implemented (ML, ANN)	Restricted application	Private universities show slower progress in analytical digital tools
Accuracy of academic performance forecasting	85–87%	80–83%	National universities maintain a slight advantage in predictive accuracy
Human resource management practices	AI-supported analytics and forecasting	Conventional HR management methods	National universities demonstrate a higher degree of HR digital transformation
Main obstacles to AI adoption	Infrastructure limitations (≈50%)	Lack of competencies and expertise (≈65%)	Private universities experience a more significant shortage of qualified personnel

Based on the analysis of existing approaches, a conceptual model for the application of artificial intelligence algorithms to support management decisions in universities has been developed. The configuration of the proposed model is illustrated in Table 3.

Table 3

Conceptual framework for the integration of artificial intelligence algorithms into university resource management processes

Model component	Description	Applied technologies / algorithms	Anticipated outcome
Input information layer	Data related to educational activities, staff performance, scientific work, and infrastructure resources	Big Data technologies, institutional databases	Creation of a comprehensive informational foundation for further analysis
Analytical module	Processing datasets, detecting trends, and revealing interconnections	Machine learning methods (ANN, SVM, DT)	Higher analytical precision and identification of latent relationships
Optimization module	Improving the allocation of schedules, workloads, and financial resources	GA, PSO, ACO, DE algorithms	Reduction of conflicts, balanced workloads, and more efficient resource distribution
Forecasting module	Predicting academic performance and personnel-related risks	Neural network technologies and ML-based models	Enhancement of prediction reliability (up to 80–87%)
Decision-support module	Assisting managerial decision-making through analytical tools	AI-driven decision support systems (DSS)	Increasing the effectiveness and justification of administrative decisions
Management outcomes	Integration of developed solutions into university governance structures	Unified digital management platforms	Growth in the overall efficiency of resource administration
Feedback mechanism	Monitoring performance indicators and refining analytical models	Data analytics and AI technologies	Ongoing improvement and adaptation of the management framework

As demonstrated in Table 3, the combined application of artificial intelligence algorithms enables the integration of analytical, predictive, and optimization functions, thereby contributing to a substantial increase in the effectiveness of university resource management processes.

Conclusion

The conducted research demonstrated the considerable opportunities associated with applying artificial intelligence algorithms to improve the management of university resources within the ongoing digitalization of higher education. The obtained results indicate that intelligent analytical and optimization tools enhance the quality of managerial decision-making and support a more efficient distribution of academic, scientific, and personnel resources.

The analysis of existing university management practices showed that although digital management systems are already widely implemented, the integration of advanced AI algorithms is still relatively limited. Nevertheless, the modeling of optimization methods such as PSO, ACO, and DE confirmed their practical

effectiveness in minimizing timetable conflicts, improving classroom utilization, balancing faculty workloads, and optimizing the allocation of research funding. In addition, the application of machine learning models significantly increased the precision of forecasting educational outcomes and identifying personnel-related risks.

Comparative assessment results revealed that national universities possess a more advanced level of digital transformation and AI integration than private universities, largely due to stronger institutional support and better-developed technological infrastructure. At the same time, several critical obstacles to broader AI implementation were identified, including underdeveloped digital infrastructure, insufficient managerial competencies, and concerns related to the security and protection of personal data.

The conceptual framework developed in this study confirms the practicality of combining analytical, predictive, and optimization instruments within a unified decision-support environment for university management. The coordinated use of such technologies promotes the transition toward a data-oriented governance model, which plays a crucial role in strengthening the competitive capacity of higher education institutions.

It should be noted that the findings presented are not universal or definitive and are limited by the sample, methodology, and research questions addressed in this article. The results reflect trends and potential applications of artificial intelligence algorithms in the context of a specific study and require further empirical testing and refinement in broader and more diverse settings.

Thus, the implementation of artificial intelligence algorithms in university resource management systems represents a strategically important area for the development of higher education. Prospects for further research lie in the development of applied solutions for integrating AI into management processes, as well as in the formation of a regulatory and ethical framework and the development of management competencies in the digital economy.

Conflict of Interest Statement

The authors declare no potential conflicts of interest regarding the research, authorship, or publication of this article.

Author Contributions

Aziza Zhunusbekova: Conceptualization, Methodology, Supervision, Proofreading. Samal Askarkyzy: Data Collection and Curation, Resources, Validation, Writing – Review and Editing. Bakhyt Alipova: Literature Review, Data Analysis, Visualization, Writing – Original Draft, Final Editing.

Declaration of Use of Generative AI and AI-Assisted Technologies

The authors declare that no artificial intelligence (AI) tools were used in the preparation of this manuscript.

References

- Astana Hub. (2026). Alem.ai: How Kazakhstan is becoming a leader in the field of artificial intelligence. <https://astanahub.com/ru/article/alem-ai-kak-kazakhstan-vykhodit-v-lidery-v-sfere-iskusstvennogo-intellekta>
- Dai, Y., Chai, C. S., Lin, P. Y., Jong, M. S. Y., Guo, Y., & Qin, J. (2020). Promoting students' well-being by developing their readiness for the artificial intelligence age. *Sustainability*, 12(16), 6597. <https://doi.org/10.3390/su12166597>
- Daniela, L. (2019). Smart pedagogy for technology-enhanced learning. In L. Daniela (Ed.), *Didactics of smart pedagogy* (pp. 3–21). Springer. https://doi.org/10.1007/978-3-030-01551-0_1
- Füller, J., Hutter, K., Wahl, J., Bilgram, V., & Tekic, Z. (2022). How AI revolutionizes innovation management: Perceptions and implementation preferences of AI-based innovators. *Technological Forecasting and Social Change*, 178, 121598. <https://doi.org/10.1016/j.techfore.2022.121598>
- Government of the Republic of Kazakhstan. (2024). Concept for the Development of Artificial Intelligence for 2024–2029 (Resolution No. 592 dated July 24, 2024). <https://adilet.zan.kz/rus/docs/P2400000592>

- Greenland, S., Saleem, M., Misra, R., & Mason, J. (2022). Sustainable management education and an empirical five-pillar model of sustainability. *The International Journal of Management Education*, 20(3), Article 100658. <https://doi.org/10.1016/j.ijme.2022.100658>
- Mahajan, R., Gupta, P., & Misra, R. (2022). Employability skills framework: A tripartite approach. *Education + Training*, 64(3), 360–379. <https://doi.org/10.1108/ET-12-2020-0367>
- Ministry of Digital Development, Innovations and Aerospace Industry of the Republic of Kazakhstan. (2025). Kazakhstan leads Central Asian countries in readiness for the implementation of artificial intelligence. <https://www.gov.kz/memleket/entities/maidd/press/news/details/1131135?lang=ru>
- Ministry of Digital Development, Innovations and Aerospace Industry of the Republic of Kazakhstan. (2026). Almost 1 million Kazakhstanis have completed training in artificial intelligence: How the instruction of the Head of State on training personnel for a digital country is being implemented. <https://www.gov.kz/memleket/entities/maidd/press/news/details/1142334?lang=ru>
- Ministry of Digital Development, Innovations and Aerospace Industry of the Republic of Kazakhstan. (2026). Kazakhstan strengthens science through artificial intelligence: Opening of the AI SilkNet center. <https://www.gov.kz/memleket/entities/maidd/press/news/details/1170236?lang=ru>
- Ministry of Education of the Republic of Kazakhstan. (2025). Kazakhstan approves national standards for the implementation of artificial intelligence in education. <https://www.gov.kz/memleket/entities/edu/press/news/details/1070018?lang=ru>
- Ministry of Science and Higher Education of the Republic of Kazakhstan. (2023). On approval of the Concept for the Development of Higher Education and Science in the Republic of Kazakhstan for 2023–2029. <https://adilet.zan.kz/rus/docs/P2300000248>
- Khatri, P., Duggal, H. K., Lim, W. M., Thomas, A., & Shiva, A. (2024). Student well-being in higher education: Scale development and validation with implications for management education. *The International Journal of Management Education*, 22(1), Article 100933. <https://doi.org/10.1016/j.ijme.2024.100933>
- National Information Technologies JSC. National artificial intelligence platform. <https://www.nitec.kz/ru/proekty/ii-platforma>
- Ng, W. C., Ravesangar, K., Toh, G. G., & Soong, Y. Q. (2025). Trends of artificial intelligence in human resources management using bibliometric analysis. In *Artificial intelligence in peace, justice, and strong institutions* (pp. xx–xx). IGI Global. <https://doi.org/10.4018/979-8-3693-9395-6.ch016>
- Ratten, V., & Jones, P. (2023). Generative artificial intelligence (ChatGPT): Implications for management educators. *The International Journal of Management Education*, 21(3), Article 100857. <https://doi.org/10.1016/j.ijme.2023.100857>
- Republic of Kazakhstan. (2025). Law of the Republic of Kazakhstan on Artificial Intelligence (No. 230-VIII ZRK, November 17, 2025). <https://adilet.zan.kz/eng/docs/Z2500000230/history>
- Perifanis, N.-A., & Kitsios, F. (2023). Investigating the influence of artificial intelligence on business value in the digital era of strategy: A literature review. *Information*, 14(2), 85. <https://doi.org/10.3390/info14020085>
- President of the Republic of Kazakhstan. (2026, January 6). On declaring the Year of Digitalization and Artificial Intelligence (Decree No. 1151). <https://www.akorda.kz/ru/ob-obyavlenii-goda-cifrovizacii-i-iskusstvennogo-intellekta-601222>
- Tambe, P., Cappelli, P., & Yakubovich, V. (2019). Artificial intelligence in human resources management: Challenges and a path forward. *California Management Review*, 61(4), 15–42. <https://doi.org/10.1177/0008125619867910>
- Xu, X. (2025). AI optimization algorithms enhance higher education management and personalized teaching through empirical analysis. *Scientific Reports*, 15, 10157. <https://doi.org/10.1038/s41598-025-94481-5>

Information about authors

Aziza Zhunusbekova - PhD, Acting Associate Professor, Department of Primary Education, Abai Kazakh National Pedagogical University, Almaty, Republic of Kazakhstan; e-mail: a.zhunusbekova@abaiuniversity.edu.kz; ORCID iD: 0000-0001-7598-4081.

Samal Askarkyzy - PhD, Acting Associate Professor, Department of Social and Humanitarian Disciplines, Kazakh National University of Arts named after K. Baiseitova, Astana, Republic of Kazakhstan; e-mail: samal050191@mail.ru; ORCID iD: 0000-0003-1325-2146 (**corresponding author**).

Bakhyt Alipova - PhD, Professor, Department of Mechanical and Aerospace Engineering, University of Kentucky, Kentucky, USA; e-mail: bakhyt.alipova-turner@uky.edu; ORCID iD: 0000-0003-0915-2759.

Received: 21.05.2026

Revised: 05.07.2026

Accepted: 28.09.2026

Published: 30.09.2026

Type of Article: Research Article

¹Askat Tleuov, ²Temirlan Tursynbayev*

¹ KIMEP University, Almaty, Republic of Kazakhstan

² Independent Agency for Accreditation and Rating, Astana, Republic of Kazakhstan

CURRICULUM INTERNATIONALISATION AND LOCAL ADAPTATION IN KAZAKHSTANI DUAL-DEGREE STEM PROGRAMMES

Abstract. Kazakhstan has expanded joint and dual degree provision; however, research has not to date examined how a local programme differs from its foreign-partner version within the same university. This study compares two STEM programme pairs at regional Kazakhstani universities: Biotechnology at Kozybayev University and its University of Arizona-linked version, and Electroenergy at Zhubanov University and its Heriot-Watt-linked version. The corpus comprises EPVO records, university programme pages, and partner-university webpages. Using qualitative comparative document analysis in an embedded multiple-case design, the study examines curriculum structure, content, learning outcomes, skills orientation, international integration, and local adaptation. The documents reveal contrasting curricular logics. Local programmes emphasize sector practice, national standards, regulation, and domestic labour-market needs. Dual-degree versions retain that base but reallocate emphasis toward industrial and research-oriented biotechnology, and toward digital, computational, sustainable, and project-based engineering. The paper contributes a programme-level account of curriculum internationalisation and names this pattern curriculum layering.

Keywords: curriculum internationalisation; dual-degree programmes; local adaptation; STEM education; document analysis.

Introduction

By the 2020/21 academic year, Kazakhstani universities had concluded almost 6,800 partnership agreements with educational institutions in 85 countries, and the number of joint and double-degree programmes developed on their basis has grown from 108 in the 2020/21 to approximately 245 by 2024 (Ministry of Science and Higher Education, as cited in The Astana Times, 2024; WENR, 2021). These figures take double-degree education beyond the scope of individual pilot projects. They describe a strategic component that the state purposefully implements from year to year in an increasing number of educational institutions.

Kazakhstan adheres to this component along with other but related commitments: full membership in the Bologna Process since 2010 and bringing its qualification structure in line with the European Higher Education Area (EHEA) (details in section 1.2). The bilateral partnerships with universities and the multilateral Bologna Agreement are based on different principles, but they are combined in the same universities, the same credit systems and often with the same students. Kazakhstan, as a post-Soviet country that joined the Bologna Process belatedly and which continues to align its national qualifications framework with European benchmarks, has reasons to use the bilateral partnership as a second, more targeted approach to implement international curriculum standards (Hartley et al., 2016). Double-degree programmes allow universities to import specific foreign curricula without waiting for system-wide harmonisation.

STEM fields are disproportionately important in this system. Kazakhstan's double degrees programme includes biotechnology, electroenergy, petroleum engineering, and information systems, disciplines directly related to its stated national priorities in agriculture, energy, and industrial diversification. A graduate of Biotechnology or an electrical engineer who was trained partly through a foreign curriculum enters the labour market, which the government has identified as strategically important.

This study investigates the following two programmes: the partnership in biotechnology between

Kozybayev University and the University of Arizona, and the partnership in electroenergy between Zhubanov University and Heriot-Watt University. In both cases, the foreign curriculum is combined with the existing national curriculum at a regional university that is not the parent university. In the documents describing these programmes, the foreign curriculum is not considered a substitute for the local one. The documents describe something closer to an addition: the national architecture of obtaining a degree is preserved, and content from foreign partners is added to it. The following chapters trace how this addition is presented in the official record, discipline by discipline, category by category.

National policy background: Kazakhstan's higher education reform trajectory

Kazakhstan acceded to the Bologna Process on 11 March 2010, at the Budapest-Vienna Ministerial Conference, becoming the process's 47th member and the first Central Asian state admitted in full standing (ENIC-Kazakhstan, n.d.; IQAA, n.d.). The date is of particular importance here because it sets the baseline: everything that is done within the framework of the two thematic study programmes, including the structure of credits, degree cycles and qualification levels, corresponds to the obligations of 2010.

The national Qualifications Framework, based on this commitment, provides for obtaining a bachelor's degree at level 6, which corresponds to the first cycle of the Qualifications Framework of the EHEA (European Training Foundation, 2021). Both electroenergy programmes examined in this paper carry that level-6 designation explicitly in their documentation.

The persistence of national degree architecture is also consistent with Kazakhstan's uneven move toward institutional autonomy. Sagintayeva and Kurakbayev (2015) locate this problem in Kazakhstan's long-standing "centralist model of institutional governance" (p. 197), where universities still have to comply with "standardized curriculum across bachelors, masters and doctoral programmes" (p. 201). Foreign-partner curriculum content therefore enters a regulated programme architecture rather than an empty institutional space.

Language policy adds a second layer to this structure. President Nazarbayev announced the trilingual education concept in October 2006 and set it out formally as "Trinity of Languages" in his 2007 policy address, assigning Kazakh the status of state language, Russian the role of interethnic communication, and English the function of global integration (Karabassova, 2020). The Ministry subsequently developed a dedicated implementation concept in 2012. STEM subjects were initially included in the English language curriculum because English has a competitive advantage that the other two languages cannot provide (Manan & Hajar, 2024). In both dual degree programmes discussed in this article, the language of instruction switches to English precisely where the trilingual policy presupposes such a transition: in science and engineering disciplines, in educational institutions beyond the flagship level. Thus, the transition to English can be considered as part of the process of spreading and implementing "global templates of knowledge" within the framework of the internationalisation of curricula (Shahjahan et al., 2024, p. 278).

This flagship tier belongs to Nazarbayev University, an autonomous educational institution with English as the language of instruction, which has become an international benchmark since its foundation. Kozybayev and Zhubanov universities take a different position. Government sources describe a policy of positioning foreign universities as strategic partners of regional educational institutions, expanding internationalisation beyond the flagship model in the state university system, which serves the majority of Kazakhstani students (WENR, 2021). On this basis, the Kozybayev-Arizona partnership was established in 2022 as the first University of Arizona microcampus in Central Asia, enrolling dual-degree biotechnology students alongside programmes in information systems and inclusive education (Arizona International, 2023; The Astana Times, 2025). The Zhubanov-Heriot-Watt partnership was established in 2023 and covers three double-diploma engineering programmes, including the Electrical and Electronic Systems degree discussed in this article (Heriot-Watt University, 2023). This detail distinguishes Zhubanov from the case of Kozybaev, where the documented model means dual certification from two named institutions, rather than a single global credential.

Conceptual background: internationalization of curriculum and transnational higher education

According to Knight (2004), internationalisation is the process of integrating an international, intercultural, or global dimension into the goals, functions, and implementation of higher education, distinguishing activities aimed at sending students and staff abroad from those that change what happens within the national curriculum itself. Leask (2015) focused on the second half of this distinction, defining the internationalisation of the curriculum as the incorporation of international and intercultural aspects into the

content, learning outcomes, assessment, and teaching methods separately from the informal curriculum concerning campus life and student interaction. The programmes discussed here fully comply with Leask's requirements. Their documents describe changes to course content, credit structures, and the language of instruction. They say almost nothing about the informal or extracurricular aspects of the student experience, which lie beyond their documentary scope and the requirements of this article.

In the narrower literature on cross-border credential arrangements, Knight (2011) and Knight and Lee (2012) separate joint degrees, which award a single combined qualification, from double or dual degrees, which award two separate qualifications in one course of study, and from consecutive degrees, which sequence two qualifications across different periods of enrolment. This typology is important for an accurate reading of the case file: the Kozybaev-Arizona programme is documented as a double degree agreement with dual certification, while the Zhubanov-Heriot-Watt programme, according to Heriot-Watt's own description, approaches a single global credential, despite having a double degree mark at the local level. McBurnie and Ziguras (2007) propose a broader category that includes both mechanisms: transnational education, defined as including any education provided by an educational institution located in one country to students living in another, whether through branch campuses, partnership agreements, or degree programmes supported by partners. Lane and Kinser's research programme on multinational colleges and universities (Lane & Kinser, 2011) and their subsequent work on cross-border privatisation (Lane & Kinser, 2011) and on the tension between global market competition and local economic development in these arrangements (Owens & Lane, 2014) provide the empirical vocabulary this paper draws on to describe partnership structure, credit transfer, and institutional governance.

Marginson and Rhoades (2002) propose a conceptual tool suitable for solving a different problem: how to consider both local and international elements simultaneously, rather than considering one as a consequence of the other. Their glonacal agency heuristic identifies three planes, global, national, and local, that operate simultaneously rather than sequentially, with agency exercised at every scale rather than flowing downward from the global to the local alone. In relation to the educational curriculum, this heuristic defies interpretation, in which the international partner simply displaces the national content. Instead, it supports a concept in which the architecture of national diplomas, national regulatory references, and the foreign partners' curricula co-exist within the same document, while maintaining their own visible presence. A recent study on the internationalisation of curricula and the development of global curricula extends this logic directly to transnational higher education. Shahjahan et al. (2024) define "content integration" and "mobility" as recurring modes of actualisation, but their more useful aspect for this paper is that the internationalisation of curricula "contextual and nuanced" and has "political and epistemological consequences" (p. 262); Caniglia et al. (2018) describe curricula that combine elements common to the whole world with implementation at the local level, rather than replacing one with the other, and Miani (2026) describes Sino-foreign joint universities in which local, national and global elements interpenetrate and create tension, rather than combining into a single coordinated project. In this document, the multilayered curriculum is used as a working term to refer to the same phenomenon in the case of Kazakhstan: the structure of the national programme is maintained as a base level, and the content of foreign partners is added on top of it, while the two layers are displayed together in the same programme document, rather than one replacing the other.

Empirical literature on joint and double-degree programmes

The empirical literature on joint and double-degree programmes focuses on outcomes rather than curriculum content. According to a survey by the Institute of International Education conducted in 245 educational institutions in 28 countries, business and management were named as disciplines for current or planned joint and double-degree programmes by 47.3% of respondents, while engineering was 39.2%, ahead of all other fields in which the survey was conducted (Obst, Kuder, & Banks, 2011). Obst and Kuder (2012) describe business management and engineering as "mainstream" of the joint-degree at the international level. Detailed documentary studies of how a particular engineering or biotechnology curriculum is assembled at a specific regional university remain rare.

In Kazakhstan, a recent study has begun to consider double-diploma programmes as independent objects of study rather than as an element of education policy. Komekbayev, Kurmangaliyeva and Kassenov (2025) analyse joint education with a double degree as a tool for improving the quality of programmes in accordance with European quality assurance standards, comparing Kazakhstan's experience with the criteria of ESG and

the EHEA. Komekbayev, Kurmangaliyeva and Dyke (2025) are studying dual-degree programmes in connection with the results of scientific research in Kazakhstan, based on the Ministry's statistical reports on joint and dual programmes for 2022 and 2023. None of the studies uses the policy documents themselves as primary material. Both studies use aggregated data at the institutional or ministerial level. Research focused on faculty shows another limitation at the participant level: Kazakhstani teachers often support internationalisation, while expressing concern about brain drain, cultural identity, and limited opportunities for mobility and research collaboration (Kuzhabekova et al., 2022). The comparative study of Pakistan and the United Kingdom conducted by Zahid and Neary (2023) is useful as a guide for neighbouring countries, as it highlights global and local tensions in employment, labour market information, and global citizenship.

The extensive literature on higher education in the post-Soviet space and Central Asia reflects the regional context in which the examples given in this article are considered, without directly addressing dual degree programmes. Chankseliani (2022) treats internationalisation as a defining feature of post-Soviet higher education transformation rather than an accidental factor. Likewise, Silova (2011) demonstrates how global reform models are accepted, adapted, and also sometimes rejected on the margins of the former Soviet Union, which bears directly on how foreign curricula are absorbed by regional Kazakhstani universities. None of this literature examines a paired local and dual-degree programme at the level of the curriculum document.

Gap statement

Two absences run through the literature reviewed above. The empirical work on joint and double-degree programmes measures outcomes: quality-assurance compliance, research output, disciplinary distribution across institutions. It does not treat the programme document itself as a site where claims about local relevance and international standard are made, balanced, and sometimes left in tension. The post-Soviet higher education literature does the opposite: it interprets policy and institutional change at a national or system level without descending to the level of a single paired programme comparison.

No study identified in this review compares a local programme and its dual-degree counterpart within the same university and the same discipline, treating the pairing itself as the unit that makes the international addition visible. Existing accounts of Kazakhstani double-diploma programmes describe them individually, against national statistics, or against European quality benchmarks. None reads a dual-degree programme beside the local programme occupying the same disciplinary space at the same institution, which is the only vantage point from which the specific curriculum content added by the foreign partner, rather than the curriculum content already present in any Kazakhstani biotechnology or electroenergy programme, becomes separable in the documentary record.

This gap is not primarily a disciplinary one. Engineering already receives more attention in the global joint-degree literature than most other fields (Obst, Kuder, & Banks, 2011). It is a gap in comparative method, and a gap specific to regional, non-flagship Kazakhstani institutions rather than to the internationalisation literature broadly, which has produced considerable work on Nazarbayev University but comparatively little on the state universities where double-diploma programmes have expanded most recently. The paired case design adopted in this paper addresses both absences directly, by fixing discipline and institution constant across the local and dual-degree comparison and by working from the programme documents as primary material rather than as background for aggregate analysis.

Purpose, research questions, and argument

This study examines how official programme documents represent differences between local and dual-degree STEM curricula within the same universities and disciplinary fields. The main research question asks: How do local and dual-degree programme documents differ in their curricular emphasis, intended graduate competencies, international integration, and adaptation to Kazakhstan's professional context?

The argument developed in the Findings section is that the programme pairs represent different but connected curricular orientations. The local programmes foreground applied sectoral preparation, national professional standards, regulatory requirements, and domestic labour-market needs. The dual-degree versions retain this foundation but redistribute curricular emphasis toward industrial and research-oriented biotechnology in one case and digital, computational, sustainable, and project-based engineering in the other. Foreign-partner content, certification, joint teaching, mobility, and project components support this reconfiguration. These claims concern the formal representation of the programmes rather than curriculum-development decisions, classroom implementation, or student experience.

Methods

This study uses a qualitative comparative analysis of documents. The design is comparative because each dual-degree programme is considered alongside a local programme in the same disciplinary field and at the same educational institution. It is document-based, as empirical material consists of official programme reports, programme descriptions, and university or partner-university webpages, rather than interviews, classroom observations, survey responses, or internal institution files.

The study is organised as an embedded project with a multiple-case design. Two institutional cases are Kozybayev University and Zhubanov University. Each case contains two programme-level units: one local programme and one dual-degree programme. The Kozybaev case involves a local Biotechnology programme and a Biotechnology programme affiliated with the University of Arizona. The Zhubanov case concerns a local Electroenergy programme and an Electroenergy programme linked to the Heriot-Watt University. Thus, the unit of analysis is the educational programme, not the university. Universities provide an institutional framework for comparison, but the analysis focuses on how the programme documentation reflects the structure of the curriculum, the language of instruction, learning outcomes, skills orientation, international integration, and adaptation to local conditions.

The aim of this study is interpretive rather than evaluative. The study does not evaluate the programme quality, teaching effectiveness, the student experience, graduate outcomes, or the implementation of partnership agreements. It examines how official documents formalise the relationship between the requirements of the local curriculum and the elements of the international curriculum. This distinction is important because programme documents may show what is specified, named, retained, added, or omitted when developing an official curriculum. They cannot show how these elements are implemented in classrooms, in assessment, student mobility, or faculty collaboration.

The cases were selected purposively. The aim was not to construct a representative sample of dual-degree programmes in Kazakhstan, but to identify programme pairs that could support a direct comparison between a local curriculum and an internationalised version in the same disciplinary and institutional setting. Four selection criteria were applied. Each case had to be located in a Kazakhstani state university; include a local programme and a dual-degree or double-diploma programme in the same broad field; involve a foreign university partner; and have publicly available programme documentation detailed enough to compare structure, language of instruction, curriculum content, learning outcomes, skills orientation, international integration, and local adaptation.

Two institutional cases met these criteria. The first is Kozybayev University, where the local 6B05102 Biotechnology programme is compared with the dual-degree 6B05111 Biotechnology programme, registered on the Unified Higher Education Platform of Kazakhstan (EPVO) under 6B05100076 and associated with the University of Arizona (EPVO, n.d.-a, n.d.-b; Arizona International, 2023). The second is Zhubanov University, where the local 6B07106 Electroenergy programme is compared with the 6B07108 Electroenergy RK+UK double-diploma programme associated with Heriot-Watt University (EPVO, n.d.-c, n.d.-d; Zhubanov University, 2025).

Table 1

Cases and programme-level units included in the study

Institutional case	Local programme unit	Dual-degree programme unit	Foreign partner	Field
Kozybayev University	6B05102 Biotechnology	6B05111 Biotechnology, EPVO registration number 6B05100076	University of Arizona	Biotechnology
Zhubanov University	6B07106 Electroenergy	6B07108 Electroenergy RK+UK	Heriot-Watt University	Electroenergy

The two cases are both STEM-oriented, but they are not identical in disciplinary content. Biotechnology and electroenergy organise knowledge, skills, and professional preparation differently. That difference is useful for this study because it allows the analysis to examine whether similar patterns of curriculum layering

appear across two technical fields, rather than inside a single programme area only.

The data corpus consists of publicly available programme-level documents. EPVO records form the core source for programme registration number, programme title, degree level, programme type, language of instruction, credit volume, learning outcomes, professional standards, and lists of courses. University and partner-university webpages were used to supplement the EPVO records where they described the foreign partnership, double-degree arrangement, mobility, joint teaching, or public-facing curriculum features. Supporting public programme databases were used only where they repeated or clarified programme information available in official sources.

The EPVO records were treated as the primary programme documents because they provide structured and comparable information across all four programmes. They identify the formal programme status, credit volume, degree level, language of instruction, professional standards, learning outcomes, and course-level information. The university and partner-university webpages were treated as supplementary documents. They were used mainly for features that EPVO does not fully describe, such as the public framing of the partnership, the foreign partner's involvement, the dual-degree arrangement, and mobility or joint-teaching claims.

Table 2

Document corpus and information extracted

Programme	Main document sources	Information extracted
6B05102 Biotechnology	EPVO record; Univision programme page	Programme status, credits, language of instruction, learning outcomes, curriculum content, skills orientation
6B05111 Biotechnology	EPVO record; Kozybayev programme page; Arizona International announcement	Double-degree status, partner university, English-medium instruction, curriculum emphasis, mobility, joint teaching, international alignment
6B07106 Electroenergy	EPVO record	Programme status, credits, language of instruction, professional standards, learning outcomes, curriculum content
6B07108 Electroenergy RK+UK	EPVO record; Zhubanov-Heriot-Watt programme page	Double-diploma status, partner university, English-medium instruction, credit volume, Heriot-Watt-linked content, professional standards, curriculum emphasis

The EPVO pages were saved as PDF files before analysis. This was necessary because EPVO is a dynamic online register and programme records may be updated over time. The saved versions provided a fixed documentary corpus for coding and comparison. The analysis therefore refers to the documents as they appeared in the saved records used for this study, rather than treating the online register as a stable text.

When preparing the documents, the study used structured extraction rather than a free summary. The study used a comparative matrix with one column for each programme and rows representing the analytical categories: programme structure and degree format, language of instruction, curriculum content, learning outcomes, skills orientation, international integration and adaptation to local conditions. The study included information in the matrix only when it was present in the documents. Absences were also recorded. For instance, if the document contains a description of the language of instruction, but there is no explanation of how this language was distributed among courses, this absence was retained as a documentary limitation and not filled in by inference.

The study reviewed the source documents in the original languages. The EPVO materials were primarily in Kazakh, the university's web pages contained materials in English and Kazakh-Russian, and information about the partner university was in English. The relevant content in Kazakh and Russian has been translated into English for analysis and manuscript writing. The translation was used as an analytical step rather than to replace the source text: programme names, learning outcomes, course descriptions, and fields indicating the language of instruction were checked against the original wording when they confirmed the findings.

Data analysis

In this study, the analysis used directed qualitative content analysis supported by a framework matrix. Since the study did not begin with open-ended coding of an unstructured corpus, the method of directed content analysis was chosen for the analysis. The research questions and document format already identified the

categories that needed comparison: degree structure, language, curriculum content, learning outcomes, skills orientation, international integration, and local adaptation. These categories guided the first cycle of coding, following the logic of directed qualitative content analysis in which prior concepts organise the initial coding structure (Hsieh & Shannon, 2005). The framework matrix then allowed the same categories to be applied across all four programme units in a systematic way (Gale et al., 2013).

The first stage was familiarisation. Each document was read to identify where it described the programme's formal structure, disciplinary content, learning outcomes, skills, language policy, partner involvement, and local or international reference points. The purpose of this reading was not to paraphrase the documents as whole texts. It was to locate comparable evidence across documents that were uneven in length, format, and level of detail.

The second stage was deductive coding. Text segments were coded under seven categories: programme structure and degree format; language of instruction; curriculum content; learning outcomes; skills orientation; international integration; and local adaptation.

The third stage was within-case comparison. Each local programme was compared with its dual-degree counterpart to identify immediate documentary differences: language shifts, foreign-partner status, curriculum additions, differences in learning outcomes, and retained local references.

The fourth stage consists of cross-case comparison. Two institutional cases were compared to identify recurring patterns and differences in specific cases. At this stage, a separation was made between what was in both cases and what belonged to one discipline or one partnership. For instance, both dual degree programmes were documented as English language learning programmes and foreign partner programmes, but the disciplinary additions were different. Biotechnology has focused on fermentation, industrial biotechnology, GMO-related content, and advanced research methods. Electroenergy has moved to computing, digital electronics, modelling, AI, sustainability, and project-based engineering.

Theme construction was based on a matrix comparison. The final topics were not identified as headings before the analysis. They were developed based on repeated patterns in the coded matrix and then refined according to the documents. This process produced the findings structure: programme configuration, national anchoring of local programmes, international curriculum layers, shifts in skills orientation and graduate profile, local adaptation within internationalised programmes, and cross-case curriculum layering. The theme-building process therefore combined deductive coding with comparative thematic interpretation, rather than treating themes as isolated topics detached from the programme pairs.

The final analytic construct was curriculum layering. In this study, curriculum layering refers to the documented addition of foreign-partner curriculum content, English-medium instruction, international certification, mobility or project components, and expanded graduate competencies onto programmes that retain national degree architecture and local professional reference points. The construct was used to name the cross-case pattern after within-case and cross-case comparisons had been completed. It was not used to claim that the programmes are implemented successfully or experienced in the same way by students and staff. The documents support claims about formal curriculum representation; they do not support claims about enactment.

Analytical consistency was addressed through a fixed comparison structure. The same categories were applied to all four programme-level units: programme structure and degree format, language of instruction, curriculum content, learning outcomes, skills orientation, international integration, and local adaptation. This reduced the risk of reading one case through a richer set of categories than the other. It also kept the comparison tied to information that appeared across the document corpus, rather than to isolated details available for only one programme.

Documentary silence was treated as part of the evidence. If the documents specify English as the language of instruction, but do not describe how English is taught, the analysis does not infer for conclusions about classroom practice. If the webpage mentioned joint teaching or mobility without specifying the details of assessment procedures, the analysis limited the relevant statement to the level of a formal description of the educational programme. This was particularly important because the study examined the official presentation of the curriculum, and not the quality of its implementation, student experience, or graduate outcomes.

The final themes were checked by paired comparison. Each topic should reflect both institutional cases, as well as the differences between the local-to-dual-degree within each case. Themes that described only one

programme, one partner, or one discipline were treated as case-specific evidence rather than cross-case findings. The resulting analysis therefore separates recurring patterns, such as English-medium dual-degree design and foreign-partner curriculum additions, from field-specific curriculum content, such as industrial biotechnology in one case and digital engineering in the other.

Ethical considerations

The present article uses publicly available programme documents. It does not report interview data, student data, personnel data, internal correspondence, unpublished institutional documents, or confidential curriculum materials. The analysed sources are official programme records, educational programme passports, university webpages, partner-university webpages, and public programme information pages. For this reason, the main ethical obligation in this article is not participant protection in the usual interview-based sense, but careful handling of institutional representation.

The universities and partner institutions are named because the programmes are publicly identifiable in official registers and university webpages. Anonymisation would not meaningfully protect the cases: the combination of programme field, foreign partner, programme code, and public dual-degree arrangement would make the institutions recognisable. The paper therefore names the programmes and institutions, but limits its claims to formal curriculum documentation. It does not evaluate institutional performance, programme quality, teaching effectiveness, or partnership success.

Findings

In this section, the findings are presented as a comparison of two local bachelor's programmes with their corresponding dual degree versions. The analysis examines each programme as it is represented in official documents. These documents define the degree structure, the language of instruction, curriculum emphasis, learning outcomes, as well as the links to the national and international professional context. However, they do not reflect the specifics of classroom teaching, student experience, graduate outcomes, or the quality of programme implementation. Therefore, the analysis is limited to the level of formal presentation of the curriculum.

The Kozybayev University pair is closely matched in structure. Both are four-year bachelor's degree programmes in biotechnology and provide for the completion of 240 credits. The local programme is described as a biotechnology programme implemented in Kazakh and Russian and developed by Kozybayev University. The dual-degree programme version retains the same length of study and credit volume, but its formal status is changing. It is presented as a double-degree programme with the University of Arizona, where the language of instruction is English. Thus, the differences are not in the basic structure of the educational programme, but rather in the presence of a partner reference, double-degree status, and the transition to English-medium instruction (EPVO, n.d.-a, n.d.-b; Univision, n.d.; Arizona International, 2023).

A pair of Zhubanov University programmes follow the same logic, but include one documented structural difference. Both programmes are four-year Bachelor's degree programmes in electroenergy and correspond to level 6 of the National Qualifications Framework. The local programme is defined as an active Kazakhstan educational programme without a foreign partner, providing 240 credits and training in Kazakh and Russian languages. The double diploma programme is defined as a new educational programme implemented jointly with Heriot-Watt University, taught in English and providing 243 ECTS credits. Differences in the credit notation a formal distinction between a local programme and a double-diploma version, while the degree level and the disciplinary field remain the same (EPVO, n.d.-c, n.d.-d; Zhubanov University, 2025).

Local educational programmes form the national disciplinary framework for pair comparisons. In official documents, the programmes are not presented as context-neutral versions of biotechnology and electroenergy. They are focused on Kazakh and Russian language instruction, sectoral application, professional standards, and nationally determined graduate training. Before the inclusion of a component related to a foreign partner in the comparison, the curriculum is already integrated into the regulatory system and labour market conditions of Kazakhstan.

The Kozybayev University biotechnology programme is described as a four-year bachelor's degree programme implemented in Kazakh and Russian. Its curriculum focuses on applied agricultural biotechnology and life sciences, including crop production, animal breeding, feed biotechnology, plant protection, biosafety,

ecology, mathematics, law, research methods, microbiology, cell biology, molecular biology, chemistry, genetics, and general biotechnological processes. The programme description is linked to the agro-biotechnological sector of Kazakhstan and the needs of the regional labour market. The learning outcomes retain the same focus: graduates are expected to know local production processes, apply quality control based on GOST standards, work in accordance with biosafety requirements, and develop managerial and communication competencies necessary for professional activity in a domestic context (EPVO, n.d.-a; Univision, n.d.).

The programme's skills profile is also close to applied production and laboratory practice. The documents focus on laboratory and field skills, biotechnological production, farming practices, and product testing. Public speaking, teamwork, and civic awareness skills are also present. Still, they are considered as elements of the learning outcomes of the local programme, rather than as international graduate characteristics. Thus, the graduate represented as a specialist in biotechnology, prepared to work in the agricultural, food, environmental and regional employment sectors of Kazakhstan, possessing technical and civil competencies formulated in a national context (EPVO, n.d.-a; Univision, n.d.).

The electroenergy programme at Zhubanov University belongs to a different sector, but shows a similar structure. It is defined as an operating Kazakhstani educational programme without a foreign partner. The programme is implemented in Kazakh and Russian, includes 240 credits and corresponds to level 6 of the National Qualifications Framework. The curriculum focuses on mathematics, physics, electrical theory, electrical machines, heat and thermal systems, power generation, as well as the design of power distribution. The programme also presents disciplines and topics focused on the national context, including General Energy, energy systems of Kazakhstan, Kazakh cultural studies and the discipline "Ulttyk Rukhaniyat" (EPVO, n.d.-c).

Practical training in the local Zhubanov University programme is described more directly in the Kozybayev University case. The document lists installation, diagnostics, maintenance, commissioning, repair, electrical measurements, power-system calculations, equipment design, testing and safe operation. The learning outcomes link these skills to local standards, national regulatory requirements, and the operation of power facilities. Transversal competencies are also presented in a national context and include teamwork, leadership, legal and ethical norms, financial literacy, environmental responsibility, Kazakh language and adherence to national regulatory requirements. In the documents, the graduate is presented as an engineer prepared to work in the energy infrastructure of Kazakhstan and in accordance with national professional standards (EPVO, n.d.-c).

The dual-degree programmes add international elements at several points in the curriculum documents. Partner names are only one part of this change. The documents also record changes in language arrangements, degree certification, curriculum content, teaching organisation, mobility options, disciplinary specialisation, and the skills attached to graduate preparation. These additions become visible because each dual-degree programme can be read beside a local programme in the same university and field.

In the Kozybayev case, the dual-degree biotechnology programme changes the programme's formal identity. The local programme is identified as a bachelor's biotechnology programme developed by Kozybayev University. The dual-degree version is identified as a Kazakhstan-US programme with dual certification involving the University of Arizona. Its language arrangement changes from Kazakh and Russian in the local programme to English in the dual-degree programme. The credit volume and four-year structure remain unchanged, but the programme is documented as an English-medium curriculum connected to a US partner and a double-degree arrangement (EPVO, n.d.-a, n.d.-b; Arizona International, 2023).

The Arizona-linked programme also reorders the disciplinary emphasis. The local biotechnology programme foregrounds agriculture-focused content: crop production, animal breeding, feed biotechnology, plant protection, biosafety, and local quality standards. The dual-degree programme foregrounds fermentation, industrial biotechnology, food and beverage technologies, biologically active preparations, insect-resistant crops, GMO risk assessment, analytical chemistry, leadership, innovation, and advanced research techniques. The programme description states that the curriculum is aligned with University of Arizona standards and prepares specialists for industry areas linked to fermented food and beverages, biologically active preparations, insect-resistant crops, and fundamental research (EPVO, n.d.-a, n.d.-b; M. Kozybayev North Kazakhstan University, n.d.; Arizona International, 2023).

The Arizona component is organisational as well as curricular. The documents mention courses developed jointly by Kozybayev University and the University of Arizona, taught with the participation of representatives of both universities, as well as the possibility of studying for one semester at the University of Arizona. These features integrate the international component into the documented structure of the programme implementation. The presented information does not show how joint teaching is conducted in classrooms or implemented in assessment procedures. Instead, they reflect a formal programme design based on the shared curriculum development, cross-institutional teaching, academic mobility and dual certification (Arizona International, 2023; M. Kozybayev North Kazakhstan University, n.d.).

The Zhubanov-Heriot-Watt programme adds a different international level. The local educational programme on electroenergy is described as an operating Kazakhstani programme with no foreign partner, implemented in Kazakh and Russian. In turn, the double diploma programme is presented as a new educational programme implemented jointly with Heriot-Watt University, taught in English and leading to diplomas from KZ and the UK. From a formal linguistic perspective, the Zhubanov-Heriot-Watt case repeats the Kozybayev-Arizona model: the local programme is documented in Kazakh and Russian, while the double diploma programme is in English (EPVO, n.d.-c, n.d.-d; Zhubanov University, 2025).

The curriculum of Heriot-Watt is clearly manifested in technical modules. The double-diploma programme retains the basic content of the electroenergy, including circuits, electronics, machines, and power systems. Then, it includes modules on advanced digital electronics, FPGA-based design, signals and systems, power system modelling, computing for engineering, sustainable energy, artificial intelligence, and senior project courses. The documents also explicitly mention modules related to Heriot-Watt, including the course “Computing for Engineering (Heriot-Watt)”. Compared to the local programme focused on power equipment, diagnostics, installation, maintenance, and Kazakhstan’s energy systems, the double-diploma version shifts the curriculum towards digital engineering, modelling, sustainable development, artificial intelligence, and project-based learning (EPVO, n.d.-c, n.d.-d; Zhubanov University, 2025).

The Zhubanov documents also link internationalisation to academic and professional skills. The double-diploma programme focuses on the development of communication and academic English, research skills, numerical methods, software tools, intercultural teamwork, inclusion and diversity, sustainability, and engineering management. The local programme pays more attention to the Kazakh language, national legislation, national ethics, legal norms, financial literacy and environmental responsibility in the Kazakhstani professional context. This distinction is descriptive rather than evaluative: one programme defines professional preparation through national industry practice, while the other complements it with an English-medium engineering profile based on the components of the Heriot-Watt curriculum and the global engineering agenda (EPVO, n.d.-c, n.d.-d; Zhubanov University, 2025).

The documents define different sets of skills for local and dual-degree programmes. Local programmes focus on applied competencies in Kazakhstan's sectoral conditions. The dual degree programmes expand a graduate's documented profile through English language proficiency, research skills, digital or advanced technical tools, project work, leadership, innovation, and international professional references.

In the Kozybayev University biotechnology case, the local educational programme represents the graduate as a specialist in applied biotechnology for agriculture, production, and laboratory work. The list of declared competencies includes laboratory and field practice, biotechnological production, farming practices, product testing, biosafety, and quality control. The learning outcomes also include knowledge of local production processes and GOST-based quality standards. Communication skills, teamwork, public speaking, and civic awareness are present in the programme description, but they are considered within the framework of local professional and social expectations, rather than through the prism of international mobility or certification by foreign partners. According to the documents, the graduate is characterised as a technically trained, practice-oriented specialist focused on the agrobiotechnological sector of Kazakhstan (EPVO, n.d.-a; Univision, n.d.).

The Biotechnology programme, implemented jointly with the University of Arizona, maintains a disciplinary foundation in microbiology, cell biology, molecular biology, chemistry, genetics, and biotechnological processes. It is complemented by two areas of competence development. The first is related to specialised industrial biotechnology and includes fermentation, food and beverage biotechnology, biologically active compounds, GMO risk assessment, analytical chemistry, and advanced research

techniques. The second area covers broader professional preparation: English language proficiency, leadership competencies, innovation management, project teamwork, data analysis, and internationally recognised industry standards. The graduate is still positioned as a biotechnology specialist, but the documented profile of his competencies is shifting towards industrial processes, research techniques, English-mediated communication, and mobility provided by partnership with the University of Arizona (EPVO, n.d.-b; M. Kozybayev North Kazakhstan University, n.d.; Arizona International, 2023).

The Zhubanov local electroenergy programme constructs a different sectoral profile. Its graduate is prepared for electrical measurements, power-system calculations, equipment design, testing, installation, commissioning, repair, diagnostics, maintenance, and safe operation of power facilities. The document also lists teamwork, leadership, legal and ethical norms, financial literacy, environmental responsibility, Kazakh language, and commitment to local regulations. Technical competence is paired with professional conduct defined through Kazakhstan's energy sector and regulatory environment (EPVO, n.d.-c).

The Heriot-Watt-linked programme keeps the electrical engineering base, including circuits, electronics, machines, and power systems, but gives more space to digital and analytical competence. The document names digital signal processing, programming, FPGA-based design, numerical methods, modelling, simulation, software tools, AI, sustainable energy, and engineering management. Senior project courses make design, creation, and problem-solving more visible in the programme description. English communication and academic study skills also enter the profile, including learning strategies, motivation, time management, reading strategies, and communication techniques. The graduate represented in this document is an electroenergy specialist whose competence is described through English-medium engineering knowledge, digital tools, project work, sustainability, and research-oriented preparation (EPVO, n.d.-d; Zhubanov University, 2025).

The dual-degree documents do not detach the programmes from Kazakhstan's educational and professional setting. Foreign-partner elements enter the curriculum, but local reference points remain: national qualification requirements, sectoral needs, domestic employment fields, regulatory knowledge, and Kazakhstan-specific professional standards. The form differs by case. In the Kozybayev documents, local adaptation appears mainly through sectoral aims, biotechnology applications, and Kazakhstan-facing regulatory references. In the Zhubanov documents, local adaptation appears through licensing, professional standards, labour-market references, and national energy priorities, even though the Heriot-Watt-linked programme is documented as English-medium.

The Arizona-linked biotechnology programme keeps several Kazakhstan-facing elements inside the internationalised curriculum. The programme addresses local needs, including insect-resistant crop development, and requires knowledge of Kazakhstan's biotechnology regulations and industries. The documents also place both biotechnology programmes within Kazakhstan's biotechnology-related employment sectors. The international additions expand the curriculum toward fermentation, industrial biotechnology, GMO-related content, and international standards, but the local sectoral reference remains present in the programme description (EPVO, n.d.-b; M. Kozybayev North Kazakhstan University, n.d.; Arizona International, 2023).

The Zhubanov-Heriot-Watt programme uses a different model. The double diploma programme is taught in English and includes content developed in collaboration with Heriot-Watt, such as computing for engineering, sustainable energy, artificial intelligence, power-system modelling, and project-based courses. Local adaptation is manifested through compliance with the licensing system of Kazakhstan, compliance with the requirements of educational legislation, national labour market needs, professional standards and consideration of national goals in the sustainable energy field. The documents indicate that both electroenergy programmes prepare graduates for professional work in Kazakhstan. Additionally, they note that the Kazakh component is less pronounced in the double diploma programme than in the local programme (EPVO, n.d.-c, n.d.-d; Zhubanov University, 2025).

The differences between Zhubanov University's two programmes are explicit. The local programme taught in Kazakh and Russian and includes the study of Kazakhstan's energy systems, national values through "Ulttik ruhaniyat," and practical preparation for local power equipment. The programme, implemented jointly with Heriot-Watt, maintains compliance with national licensing requirements and labour market needs, but focuses more on English, global engineering content, sustainability, AI, digital systems, and international

professional preparation. The comparison also shows that Kazakh and Russian language preparation is reflected in the documentation of the double diploma programme to a lesser extent than in the documentation of the local programme (EPVO, n.d.-c, n.d.-d; Zhubanov University, 2025).

Cross-case pattern: curriculum layering across both cases

The cross-case pattern is addition, redistribution, and selective retention. The dual degree programmes are not presented in the documents as a substitute for local educational programmes. The basic bachelor's architecture remains visible: in both cases, the programmes retain their affiliation to the same subject areas, the four-year structure and compliance with the requirements of the education regulation of Kazakhstan. The changes relate to those components that are being added to this structure: English instruction, partner certification, specialised disciplinary content, academic and professional skills, mobility or project components, as well as international reference points.

In both cases, four levels are repeated. The first is linguistic: local programmes are implemented in Kazakh and Russian, while dual degree programmes are documented as English language programmes. The second is disciplinary: the biotechnology programme is expanding with content related to fermentation, industrial biotechnology, and GMO-related content, while the electroenergy programme is complemented by computing, AI, digital electronics, modelling, and sustainability. The third level is organisational: dual certification, joint teaching, academic mobility, Heriot-Watt modules and curriculum elements based on the UK model link the programme with a foreign partner university. The fourth level is professional: international standards, research skills, project work, leadership, innovation, English communication, as well as intercultural or sustainability-oriented competencies are added to the graduate's profile, which is still focused on the national labour market.

The local component is unevenly preserved in the programmes. In the case of the Kozybayev-Arizona programme, local and international elements are combined through Kazakhstan-oriented biotechnology applications, links to the local industry, and regulatory knowledge, rather than by preserving the Kazakh/Russian-language profile of the programme. In the case of Zhubanov and Heriot-Watt universities, English-language instruction and the content developed jointly with Heriot-Watt form a clearer distinction between the local programme and the double diploma programmes. The documents of the Zhubanov University continue to associate the double diploma programme with the licensing system of Kazakhstan and the needs of the national labour market, but also note a less pronounced attention to disciplines focused on Kazakhstani specifics, as well as the possible need for additional language support in Kazakh and Russian. (EPVO, n.d.-b, n.d.-d; M. Kozybayev North Kazakhstan University, n.d.; Zhubanov University, 2025).

The finding is quite specific: dual degree programmes preserve the national architecture of educational programmes and their connection to the local professional context, while simultaneously supplementing this framework with the curricula of a foreign partner university, English-medium instruction, international certification, and expanded graduate competencies.

Discussion

This study aimed to identify how the official documents of educational programmes reflect the combination of international curriculum with national regulatory requirements, linguistic and labour market needs of Kazakhstan in two dual-degree STEM programmes. The answer is neither curriculum replacement nor simple curriculum transfer. In both cases, the documents considered to preserve the basic architecture of the Kazakhstan bachelor's programme and complement it with content developed by a foreign partner university, English-medium instruction, partner certification, components of academic mobility or project activities, as well as an expanded set of graduate competencies. A pair comparison makes this pattern visible. Each dual-degree programme could be interpreted as an international partnership. However, a comparison with the corresponding local programme makes it possible to determine what elements have been added, what remains, and what becomes less visible.

The findings locate internationalisation inside the formal curriculum record. This is where Leask's (2015) distinction becomes useful. The documents do not give much evidence about informal student experience, intercultural campus life, or everyday teaching. They name courses, outcomes, credits, language of instruction, partner involvement, professional standards, and intended graduate competencies. Curriculum internationalisation is therefore represented through what the programme document lists and reorganises.

Shahjahan et al. (2024) identify content integration and mobility as recurrent ways of actualising curriculum internationalisation. The Kozybayev-Arizona programme records a semester exchange option and joint teaching, yet its larger documentary change is curricular, in the shift toward industrial biotechnology and English-mediated professional preparation. The Zhubanov-Heriot-Watt programme names a stronger project-based partner-curriculum structure, and its most visible changes are likewise curricular. Mobility appears, but the curriculum document itself is the main site where internationalisation is made legible.

For Kazakhstan, this difference shifts the focus from a simple counting of international partnerships to a reading of the changes in the curriculum that these partnerships bring. In policy and public reports, the presence of foreign branches, double diploma programmes, and academic mobility of students is often seen as an indicator of internationalisation. However, the documents examined in this study point to a more specific mechanism: internationalisation is included in the structure of the educational programme through the redistribution of disciplinary content, language certification, and graduate attributes. It does not remain an external sign of international partnership in relation to the curriculum. It also does not displace the local educational programme. Thus, in the official documentation, internationalisation is presented as a layered curricular arrangement.

The content of the curriculum developed by a foreign partner university is recorded in the official documentation of educational programmes administered in the national system. Dual degree programmes remain bachelor's degree programmes embedded in Kazakhstan's qualification structures, credit systems, professional standards, licensing references, and employment fields. This finding empirically confirms the concept of the glonacal reading of higher education proposed by Marginson and Rhoades (2002), according to which the global, national and local levels function simultaneously, rather than in a single downward sequence. A foreign partner university does not stand outside the national system. Its curriculum is registered through national categories of educational programmes.

The analysis of Kazakhstan's limited transition to institutional autonomy presented by Sagintayeva and Kurakbayev (2015) helps explain the same pattern. The national architecture remains clearly represented in the documents, as dual-degree programmes are created in a system where universities may develop new partnerships and curricula, but registration of educational programmes, the level of degree awarded, educational standards and official documentation are still under significant regulatory influence. Therefore, the finding of the study is not that internationalisation weakens the national curriculum. On the contrary, the documents show that it is embedded in it.

The international curriculum layer is not generic. In biotechnology, it is built through industrial fermentation, food and beverage technologies, biologically active preparations, insect-resistant crops, GMO risk assessment, analytical chemistry, research techniques, leadership, and innovation. In electroenergy, it is built through computing, digital electronics, FPGA design, modelling, AI, sustainable energy, senior projects, engineering management, and academic study skills. The partner universities do not add the same international vocabulary to both programmes. Each field receives a different international layer.

That disciplinary difference is analytically useful for the study of STEM curriculum internationalisation. Shahjahan et al. (2024) caution that curriculum internationalisation takes shape through disciplinary and regional conditions. The present cases support that claim at programme-document level. Biotechnology is internationalised through industrial and research-oriented life-science content linked to agriculture, food, biosafety, and biotechnology industries. Electroenergy is internationalised through digital engineering, modelling, sustainability, and project-based design. These additions fit the internal logic of each field. They also show why programme-level comparison is needed. A national count of dual-degree programmes cannot show which content travels into which discipline, or which local references remain attached after the international layer is added.

The STEM setting also creates a certain epistemological tension. The technical content presented in the documents is often described as universal and portable. These elements include AI, modelling, fermentation, analytical chemistry, sustainable energy, and advanced research techniques. Local adaptation manifests itself through application, regulations, employment sectors, professional standards, and national infrastructure. At the same time, the documents do not consider Kazakhstani and international knowledge as equivalent categories. The international component is mainly related to advanced technical, digital, research, and English-medium terms, while the local component is related to regulatory, sectoral, linguistic, and

employment terms. However, such a structure should not be interpreted as a judgement of the quality of one of these components. This is a documented distribution of authority and functions.

The stable signs of local markers remain outside the language profile of educational programmes. Both dual degree programmes are switching to English-medium instruction, and this transition increases the distance from the Kazakh and Russian language basis of local programmes. However, the documents on the dual degree programmes do not separate the programmes from Kazakhstan. The Arizona-linked programme maintains links with biotechnological applications focused on the needs of the domestic sector, including insect-resistant crops, biologically active preparations, industry references, and knowledge of Kazakhstan's biotechnology regulations. The Heriot-Watt-linked programme keeps national licensing system, professional standards, labour market needs and national energy priorities.

This asymmetry can be seen clearly. Local significance is still present in the documents, but it is expressed primarily through sectoral and regulatory components, rather than through the language of instruction. This is especially noticeable in Zhubanov University programmes, where the Kazakh-specific content, training in Kazakh and Russian, Kazakhstan's energy systems and national values are presented more fully in the local programme than in the double-diploma programme. The Heriot-Watt-linked programme retains its stated focus on the needs of the national labour market, but the curriculum content is more organised around English, digital systems, AI, sustainability, project work, and UK-linked engineering components.

This does not mean that local adaptation is poorly implemented in practice. The presented data cannot support that conclusion. They only indicate that local adaptation is reflected unevenly in the official documentation. Some of its elements are clearly represented by standards, licensing, regulations, and employment fields. Others are less well covered: language support, assessment mediation, professional terminology in Kazakh and Russian, and the relation between English-medium technical content and local workplace communication. Such documentary omission is one of the findings of the study. The documents identify English as the language of instruction, but do not explain how students transition from English-medium academic content to professional settings in Kazakhstan.

This study complements the existing work on the internationalisation of higher education in Kazakhstan. Existing studies have focused on issues of institutional autonomy, system reform, internationalisation of scientific activity, quality assurance alignment, faculty perception, as well as general trends in the development of dual degree programmes. These studies help to understand the policy and institutional setting, but only rarely reveal the internal curriculum arrangements through which international components become part of a named educational programme.

The main methodological contribution of the research is the paired design. While keeping the institute and broad discipline constant, it separates the additions offered by foreign partners from the content already present in local biotechnology and electroenergy programmes. Without the local counterpart, elements such as fermentation, AI, sustainability, project work, or English-medium could be perceived as a common design of an educational programme. A paired comparison shows them as added components, redistributions, and reweightings. In addition, this approach prevents considering the local programme only as a passive background for analysis. The local programme provides a baseline for the national sector that makes the international level clear.

For curriculum internationalisation scholarship, this study makes a modest conceptual gain. Curriculum layering refers to a documented structure in which global, national and local levels are simultaneously represented in an educational programme record. This term is useful since it avoids two weak interpretations: the first, according to which the foreign partner completely transfers its curriculum to Kazakhstan, and the second, according to which local adaptation absorbs the foreign curriculum, so that its international component fully disappears. The documents show neither. They show layered programmes with retained national architecture, added foreign-partner content, English-medium delivery, partner certification, mobility or project components, and local professional references.

Conclusion

This study examined how official documents represent the relationship between local and international curriculum elements in two Kazakhstani dual-degree STEM programme pairs. The documents present curriculum internationalisation as layering: the national programme architecture remains visible, while

foreign-partner content and forms of academic organisation are added to it.

The local and dual-degree programmes therefore represent different but connected curricular orientations. The local programmes foreground applied preparation for Kazakhstan's biotechnology and energy sectors, national professional standards, regulation, and domestic employment needs. The dual-degree versions retain these foundations but redistribute disciplinary content toward industrial, research-oriented, digital, computational, sustainable, and project-based preparation. Local adaptation remains selective: it appears through sectoral applications and biotechnology regulation in one case, and through licensing, professional standards, employment fields, and energy-sector priorities in the other. Curriculum layering names this combination of retained national foundations, redistributed disciplinary content, foreign-partner organisation, and expanded graduate competencies.

Implications for policy and institutional practice

The first implication concerns the national register of educational programmes. The Register already functions as the public documentary site where programme title, credit volume, language of instruction, learning outcomes, professional standards, and programme status become publicly available. The Rules for maintaining the Register define it as the official mechanism for listing educational programmes implemented by higher and postgraduate education organisations (Minister of Science and Higher Education of the Republic of Kazakhstan, 2022). For dual-degree and joint programmes, this register could require a short curriculum-internationalisation mapping field. That field should not be a promotional description. It should ask for four items: the foreign-partner curriculum elements added to the local programme base; the Kazakhstan-specific standards, sectoral references, and professional requirements retained in the programme; the language model, including any planned support for Kazakh and Russian professional terminology; and the certification, mobility, joint-teaching, or assessment arrangements attached to the foreign partner. Such a requirement would not create a new quality label. It would make the curriculum relation more transparent for students, employers, regulators, and researchers. It would also fit Kazakhstan's current policy direction, since the Concept for the Development of Higher Education and Science for 2023-2029 names internationalisation, joint and double-degree programmes, and partnerships with foreign universities as part of the sector's development agenda (Government of the Republic of Kazakhstan, 2023).

The second implication concerns programme teams at Kazakhstani universities. English-medium dual-degree STEM programmes should contain an explicit local-adaptation note in the programme documentation. This note should explain how the programme connects English-medium disciplinary content to Kazakhstan's professional standards, local regulatory vocabulary, workplace communication needs, internships or applied projects, and sectoral employment contexts. The need is clearest in programmes where Kazakh and Russian disappear from the language-of-instruction field but the graduate is still expected to work in Kazakhstan. The requirement can be written at course-team level. A biotechnology programme could specify how students encounter Kazakhstan's biosafety, agro-biotechnology, product-quality, and regulatory terminology across English-medium courses. Similarly, the electroenergy programme can specify in which disciplines students work with Kazakhstan's energy legislation, grid terminology, safety rules, technical documentation, and employer-facing communication in Kazakh or Russian. The aim is not to reduce the international English-medium component of the programme. It is to document the relationship between this layer and the national professional environment in which many graduates are likely to move.

Limitations

The first limitation is the corpus of documents used. The study is based on the public EPVO records, university educational programmes pages and the websites of foreign partner universities. The study does not include internal curriculum maps, syllabi, assessment assignments, moderation records, academic mobility agreements, faculty workload documents, teaching materials, student work, and employer feedback. Therefore, the study can show how educational programmes are presented in the official documentation. It cannot show how the teaching of courses is implemented, how assessment standards are applied, how students perceive learning in English, and how teaching is coordinated between foreign and local faculty.

The second limitation is the case scope. The study considers pairs of two STEM programmes implemented at two regional universities in Kazakhstan. This design allows for a detailed comparison of biotechnology and electroenergy programmes, but does not reflect the full diversity of dual-degree programmes in Kazakhstan. Educational programmes in medicine, pedagogy, IT, business, social sciences and

humanities may combine international and local components in other ways. In addition, leading national and private universities may demonstrate different models of documenting such programmes. Therefore, the results obtained should be considered as analytically transferable and not as statistically generalisable.

Future research could be conducted within the same programme pairs framework that goes beyond publicly available documents. The next direction may be to study the chain of curriculum implementation from registration in the EPVO to syllabus, assignment brief, assessment rubric, teaching language, mobility record, and student work. Such a research design will make it possible to determine whether the programme levels identified in this article remain separate in the teaching process, merge through assessment, or generate tensions for students and faculty. In addition, it will help to directly explore the language aspect: how students of English-medium STEM programmes of a dual-degree master the professional terminology in Kazakh and Russian necessary to work in the national professional space, and at what stages of the implemented curriculum this learning is reflected.

Funding Information

This work was supported by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. AP23488831).

Conflict of Interest Statement

The authors declare that they have no conflicts of interest with respect to the research, authorship, or publication of this article.

Author Contributions

Askat Tleuov led the conceptualisation and methodological design of the study, supervised the data collection, wrote the original manuscript, interpreted the findings, and developed the study's conclusions and implications. Temirlan Tursynbayev collected and organised the documentary data, developed the comparison matrix, conducted the comparative analysis, reviewed and revised the manuscript, and formatted it for publication. As the corresponding author, he also managed the journal submission. Both authors reviewed and approved the final version of the manuscript.

Declaration of Use of Generative AI and AI-Assisted Technologies

The authors declare that no artificial intelligence (AI) tools were used in the preparation of this manuscript.

References

- Arizona International. (2023, January 23). UArizona partners with Kozybayev University in Northern Kazakhstan to launch our newest microcampus and the first in Central Asia. University of Arizona. <https://international.arizona.edu/news/uarizona-partners-kozybayev-university-northern-kazakhstan-launch-our-newest-microcampus-and>
- Caniglia, G., John, B., Bellina, L., Lang, D. J., Wiek, A., Cohmer, S., & Laubichler, M. D. (2018). The glocal curriculum: A model for transnational collaboration in higher education for sustainable development. *Journal of Cleaner Production*, 171, 368–376. <https://doi.org/10.1016/j.jclepro.2017.12.207>
- Chankseiani, M. (2022). *What happened to the Soviet university?* Oxford University Press.
- EPVO. (n.d.-a). 6B05102 Биотехнология. Register of Educational Programmes. Retrieved July 7, 2026, from https://epvo.kz/#/register/education_program/application/62187
- EPVO. (n.d.-b). 6B05111 Биотехнология. Register of Educational Programmes. Retrieved July 7, 2026, from https://epvo.kz/#/register/education_program/application/52657
- EPVO. (n.d.-c). 6B07106 Электроэнергетика. Register of Educational Programmes. Retrieved July 7, 2026, from https://epvo.kz/#/register/education_program/application/64086
- EPVO. (n.d.-d). 6B07108 Электроэнергетика (ҚР + UK). Register of Educational Programmes. Retrieved July 7, 2026, from https://epvo.kz/#/register/education_program/application/69757
- European Training Foundation. (2021). *National qualifications framework: Kazakhstan*. <https://www.etf.europa.eu>
- Gale, N. K., Heath, G., Cameron, E., Rashid, S., & Redwood, S. (2013). Using the framework method for the

- analysis of qualitative data in multi-disciplinary health research. *BMC Medical Research Methodology*, 13, Article 117. <https://doi.org/10.1186/1471-2288-13-117>
- Government of the Republic of Kazakhstan. (2023, March 28). Об утверждении Концепции развития высшего образования и науки в Республике Казахстан на 2023–2029 годы [On approval of the Concept for the Development of Higher Education and Science in the Republic of Kazakhstan for 2023–2029] (Resolution No. 248; as amended by Resolution No. 471 of June 14, 2024). Adilet Legal Information System. <https://adilet.zan.kz/rus/docs/P2300000248>
- Hartley, M., Gopaul, B., Sagintayeva, A., & Apergenova, R. (2016). Learning autonomy: Higher education reform in Kazakhstan. *Higher Education*, 72(3), 277–289. <https://doi.org/10.1007/s10734-015-9953-z>
- Heriot-Watt University. (2023). Heriot-Watt forms pioneering strategic partnership in Kazakhstan. <https://www.hw.ac.uk/news/2023/heriot-watt-forms-pioneering-strategic-partnership-in-kazakhstan>
- Higher Education Development National Center [ENIC-Kazakhstan]. (n.d.). History of Bologna process. https://enic-kazakhstan.edu.kz/en/bologna_process/history
- Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277–1288. <https://doi.org/10.1177/1049732305276687>
- Independent Agency for Quality Assurance in Education [IQAA]. (n.d.). Kazakhstan joined the Bologna Process. <https://iqaa.org>
- Karabassova, L. (2020). Understanding trilingual education reform in Kazakhstan: Why is it stalled? In D. Egéa (Ed.), *Education in Central Asia (Education, Equity, Economy, Vol. 8, pp. 37–51)*. Springer. https://doi.org/10.1007/978-3-030-50127-3_3
- Knight, J. (2004). Internationalization remodeled: Definition, approaches, and rationales. *Journal of Studies in International Education*, 8(1), 5–31. <https://doi.org/10.1177/1028315303260832>
- Knight, J. (2011). Doubts and dilemmas with double degree programs. *RUSC: Universities and Knowledge Society Journal*, 8(2), 297–312. <https://doi.org/10.7238/rusc.v8i2.1067>
- Knight, J., & Lee, J. (2012). International joint, double, and consecutive degree programs: New developments, issues, and challenges. In D. K. Deardorff, H. de Wit, J. D. Heyl, & T. Adams (Eds.), *The SAGE handbook of international higher education*. SAGE.
- Komekbayev, Y., Kurmangaliyeva, D., & Dyke, G. J. (2025). Unlocking scientific potential through dual degrees in Kazakhstan's higher education system. *Iberoamerican Journal of Science Measurement and Communication*, 5(2), 1–17. <https://doi.org/10.47909/ijsmc.225>
- Komekbayev, Y., Kurmangaliyeva, D., & Kassenov, K. (2025). Joint double degree education as a new standard of educational program quality. *Higher School of Kazakhstan*, 52(4), 36–49. <https://doi.org/10.59787/2413-5488-2025-52-4-36-49>
- Kuzhabekova, A., Bilyalov, D., & Mussabayeva, M. (2022). Faculty views on internationalization in Kazakhstan. *Yükseköğretim Dergisi*, 12(Supplement), 35–44. <https://doi.org/10.2399/yod.21.202204>
- Lane, J. E., & Kinser, K. (2011). Reconsidering privatization in cross-border engagements: The sometimes public nature of private activity. *Higher Education Policy*, 24(2), 255–273. <https://doi.org/10.1057/hep.2011.2>
- Leask, B. (2015). *Internationalizing the curriculum*. Routledge. <https://doi.org/10.4324/9781315716954>
- M. Kozybayev North Kazakhstan University. (n.d.). 6B05103 Biotechnology. <https://ku.edu.kz/en/educationalProgram/passport?id=1144>
- Manan, S. A., & Hajar, A. (2024). Understanding English medium instruction policy from the perspectives of STEM content teachers in Kazakhstan. *TESOL Journal*. <https://doi.org/10.1002/tesj.847>
- Marginson, S., & Rhoades, G. (2002). Beyond national states, markets, and systems of higher education: A glonacal agency heuristic. *Higher Education*, 43(3), 281–309. <https://doi.org/10.1023/A:1014699605875>
- McBurnie, G., & Ziguras, C. (2007). *Transnational education: Issues and trends in offshore higher education*. Routledge. <https://doi.org/10.4324/9780203968352>
- Miani, M. (2026). Glocal tensions in transnational higher education in China: The case of a foundation year in a Sino-foreign cooperative university. *Higher Education Quarterly*. <https://doi.org/10.1111/hequ.70085>
- Minister of Science and Higher Education of the Republic of Kazakhstan. (2022, October 12). Об

- утверждении Правил ведения реестра образовательных программ, реализуемых организациями высшего и (или) послевузовского образования, а также основания включения в реестр образовательных программ и исключения из него [On approval of the Rules for maintaining the register of educational programmes implemented by organisations of higher and/or postgraduate education, and grounds for inclusion in and exclusion from the register] (Order No. 106). Adilet Legal Information System. <https://adilet.zan.kz/rus/docs/V2200030139>
- Obst, D., & Kuder, M. (2012). International joint- and double-degree programs. *International Higher Education*, 66, 5–7.
- Obst, D., Kuder, M., & Banks, C. (2011). Joint and double degree programs in the global context: Report on an international survey. Institute of International Education. <https://www.iie.org/wp-content/uploads/2022/12/Joint-Double-Degree-Survey-Report-2011.pdf>
- Owens, T. L., & Lane, J. E. (2014). Cross-border higher education: Global and local tensions within competition and economic development. *New Directions for Higher Education*, 2014(168), 69–82. <https://doi.org/10.1002/he.20114>
- Sagintayeva, A., & Kurakbayev, K. (2015). Understanding the transition of public universities to institutional autonomy in Kazakhstan. *European Journal of Higher Education*, 5(2), 197–210. <https://doi.org/10.1080/21568235.2014.967794>
- Shahjahan, R. A., Miao, S., & Baizhanov, S. (2024). Actualizing curriculum internationalization: An integrative review. *Comparative Education Review*, 68(2), 262–285. <https://doi.org/10.1086/729664>
- Silova, I. (Ed.). (2011). *Globalization on the margins: Education and post-socialist transformations in Central Asia*. Information Age Publishing.
- The Astana Times. (2024, July). Kazakhstan bolsters foreign university partnerships, promotes academic research. <https://astanatimes.com/2024/07/kazakhstan-bolsters-foreign-university-partnerships-promotes-academic-research/>
- The Astana Times. (2025, August). Kozybayev University students study in US under dual-degree program. <https://astanatimes.com/2025/08/kozybayev-university-students-study-in-us-under-dual-degree-program/>
- Univision. (n.d.). 6B05102 Биотехнология в СКУ им. М. Козыбаева. Retrieved July 7, 2026, from <https://univision.kz/edu-program/62187.html>
- WENR. (2021, July). Education in Kazakhstan. *World Education News + Reviews*. <https://wenr.wes.org/2021/07/education-in-kazakhstan>
- Zahid, G., & Neary, S. (2023). Faculty members' perceptions of internationalization of curriculum: Globalization and localization, a comparative study. *Asian Education and Development Studies*, 12(2–3), 166–180. <https://doi.org/10.1108/AEDS-03-2023-0024>
- Zhubanov University. (2025, January 16). Electrical and electronic systems. https://zhubanov.edu.kz/en/campus_hw/category/14/

Information about authors

Askat Tleuov – PhD in Education, Associate Professor, Research Director, College of Human Sciences and Education, KIMEP University, Almaty, Republic of Kazakhstan; e-mail: a.tleuov@kimep.kz; ORCID iD: 0000-0001-6642-3047

Temirlan Tursynbayev – MSc in Educational Leadership, Project Manager for International Cooperation, Independent Agency for Accreditation and Rating, Astana, Republic of Kazakhstan; e-mail: temirlan.tursynbayev@alumni.nu.edu.kz; ORCID iD: 0009-0008-7791-1642 (*corresponding author*).

Received: 23.07.2026

Revised: 25.09.2026

Accepted: 27.09.2026

Published: 30.09.2026

Type of Article: Research Article