DOI: https://doi.org/10.37547/mesmj-V7-I4-57 Pages: 513-522

MENTAL ENLIGHTENMENT SCIENTIFIC – METHODOLOGICAL JOURNAL

MENTAL ENLIGHTENMENT SCIENTIFIC – METHODOLOGICAL JOURNAL

http://mentaljournal-jspu.uz/index.php/mesmj/index

THEORETICAL AND METHODOLOGICAL FOUNDATIONS OF ASSESSING SCHOOL STUDENTS' ACADEMIC ACHIEVEMENTS BASED ON ARTIFICIAL INTELLIGENCE

Sherzod Mamasharipovich Pardayev
Senior Lecturer
Jizzakh State Pedagogical University
sherzodpardayev84@gmail.com
Jizzakh, Uzbekistan

ABOUT ARTICLE

Key words: artificial intelligence, digital education, assessment of academic achievements, automated testing, individualized education, assessment criteria, quality improvement in education, forecasting, digital learning environment, innovative pedagogical technologies..

Received: 01.05.26
Accepted: 02.05.26
Published: 03.05.26

Abstract: This study analyzes the theoretical and methodological foundations of assessing school students' academic achievements based on artificial intelligence. In the context of the modern digital educational environment, improving the assessment process and enhancing its accuracy, transparency, and efficiency is considered a pressing issue. From this perspective, the study explores the potential of artificial intelligence technologies in the educational process and the ways of integrating them into the assessment system. The work scientifically substantiates approaches aimed at the comprehensive assessment of students' individual characteristics, cognitive activity, and learning outcomes. Furthermore, the principles, criteria, and indicators for developing assessment models based on artificial intelligence are defined. Within the framework of the study, mechanisms of adaptive assessment, automated testing, forecasting, and personalized learning support are analyzed. As a result, an innovative model for assessing students' academic achievements in school education is proposed,

and its theoretical and practical significance is justified. This approach serves to improve the quality of education, optimize the learning process, and shape the individual development trajectory of students.

Introduction. The assessment process in education has evolved in close connection with the history of human civilization. In ancient China, examinations were used as early as 2000 BCE to select civil servants, while in ancient Greece, thinkers such as Socrates, Plato, and Aristotle assessed their students' knowledge through oral dialogues. During the Middle Ages, with the establishment of the first universities in Europe, assessment systems began to take on a formal character.

The theoretical and methodological foundations of assessment began to form in the late 19th and early 20th centuries with the development of psychology and pedagogy. During this period, the first standardized tests were created by scholars such as Alfred Binet, Theodore Simon, and Edward Thorndike. In the mid-20th century, Benjamin Bloom's taxonomy played a significant role in systematizing the goals and criteria of assessment. [2, 3]

In Uzbekistan, the assessment system primarily operated on a five-point scale. Although this system was distinguished by its simplicity in assessing students' knowledge, it had shortcomings such as subjectivity and an inability to fully reflect differences. During the years of independence, fundamental reforms were implemented in the education system; notably, from 2012, general secondary schools transitioned to a 100-point assessment system. [1]

Currently, assessment systems in education are developing in the following main directions:

Formative and Summative Assessment. Formative assessment focuses on monitoring students' mastery during the learning process, providing them with rapid feedback, and correcting the learning process. Summative assessment serves to determine the level of students' mastery at the end of a particular stage. Artificial intelligence technologies create significant opportunities, especially in formative assessment, as they enable real-time analysis and the provision of personalized recommendations.

Criterion-Referenced and Norm-Referenced Assessment. In criterion-referenced assessment, a student's results are compared against predetermined criteria. In norm-referenced assessment, a student's results are compared against the results of other students in the group. Artificial intelligence systems allow for the implementation of both approaches; however, in criterion-referenced assessment, it is easier to define specific criteria and adhere to them.

Computer-Based Tests. With the development of computer technology in the late 20th century, computer-based tests became widespread. Initially, these tests were electronic versions of traditional paper tests; later, they evolved into more complex systems that utilized multimedia capabilities and included interactive tasks. Today, computer-based tests exist in complex forms such as adaptive tests, gamified tests, and simulations. [2,7,10]

Online Tests and Challenges. The development of internet technologies led to the widespread adoption of online tests. As noted in the research by S. Abdumominova and A. Akbaraliyeva, while online tests offer advantages such as economic efficiency, scalability, and rapid feedback, they also present several challenges: academic dishonesty, technical reliability, unequal accessibility, and security issues. Technologies such as remote proctoring, biometric verification, and AI-driven behavior analysis are being used to prevent academic dishonesty in online tests. [4, 5]

Adaptive Testing Systems. Adaptive testing systems adjust the difficulty level of test questions based on the student's responses. The theoretical foundations of these tests were developed by scholars such as G. Rasch, F. Lord, and D. Weiss. Compared to traditional tests, adaptive tests allow for a more accurate determination of a student's knowledge level in a shorter period. Artificial intelligence technologies further expand the capabilities of adaptive tests, as they can analyze students' responses not only as correct/incorrect but also semantically. [2, 3,4]

Gamified Assessment. In recent years, there has been a growing trend of gamification in education. In gamified assessment, students accumulate points, ratings, and badges by completing various tasks. This increases student motivation and enhances interest in the learning process. Tasks such as mosaics and crosswords are highly suitable for gamified assessment.

AI-Based Assessment. The rapid development of artificial intelligence technologies in the second decade of the 21st century has created new possibilities for assessment systems. AI-based assessment systems differ from traditional systems in the following aspects: assessing not only the final result but also the process; analyzing student actions in real-time; ensuring an individualized approach; assessing open-ended responses based on semantic analysis; identifying the causes of errors and forecasting; and providing personalized recommendations.

AI-based assessment systems represent a qualitatively new stage in the educational process. The research reveals that assessment systems have undergone the following stages in historical development: traditional assessment, computer-based tests, online tests, adaptive

tests, gamified assessment, and AI-based assessment. Each stage is characterized by its own technological solutions and methodological approaches.

An analysis of the potential of artificial intelligence technologies in education shows that they enable not only the automation of the assessment process but also an in-depth analysis of students' learning processes, the formation of individual learning trajectories, and the development of cognitive and metacognitive skills. In particular, technologies such as computer vision, large language models, agent-based systems, and semantic analysis create new opportunities in the assessment process.

Analysis of foreign experience (using the examples of the USA, South Korea, China, and European countries) revealed the following main trends in the implementation of AI-based assessment systems: the widespread use of adaptive learning platforms, real-time monitoring and analysis capabilities, the formation of personalized learning paths, and the development of systems aimed at fostering metacognitive skills. Advanced examples in this area include South Korea's "AI tutor" system, adaptive platforms from the USA such as Knewton and DreamBox, and China's Squirrel AI system.

Research by I. Okhunov and K. Mamadaliyeva highlights that artificial intelligence enables the automation of assessment, ensures transparency, enhances accuracy, and creates opportunities for analyzing student activity in real-time.

Metacognitive Monitoring. In the research by Z. Khalikova, metacognitive monitoring is described as students' ability to monitor their own knowledge, self-assess, and consciously manage the learning process. The use of AI-based interactive programs, virtual assistants, and adaptive systems helps develop metacognitive monitoring in students. Tasks such as crosswords and mosaics play a crucial role in developing metacognitive monitoring because they require students to plan, control, and evaluate their own actions. [6, 8,9]

The development trends of modern assessment systems can be summarized in Table 1.1 below:

Table 1.1. Stages of Development of Assessment Systems

Stage	**Main Characteristics**	**Period**	**Technologies**
Traditional assessment	Until the end of the 20th century	Oral and written exams, 5–10 point grading systems	Paper, pen
Computer-based tests	Test automation, fast calculation of results	1980–2000	Personal computers, simple software
Online tests	Conducting tests via the Internet,	2000–2010	Internet, web servers

	centralized databases		
Adaptive tests	Adapting to student level, Item Response Theory	2010–2015	Statistical models, algorithms
Gamified assessment	Increasing motivation, interactive tasks	2015–2020	Multimedia, mobile applications
AI-based assessment	Process analysis, semantic assessment, prediction	2020–present	Machine learning, neural networks, LLM

In recent years, significant work has been carried out in Uzbekistan regarding the improvement of assessment systems. This is evidenced by the growing body of scientific research, including the dissertation by K. K. Seitnazarov on "Improving the Methodology of Assessing Students' Knowledge in Higher Education Institutions Based on Artificial Intelligence," the research by A. B. G'oyibnazarova on assessment methodology using adaptive test systems in teaching biology, and the studies by I. Ovkhunov and K. Mamadaliyeva on automating the assessment process in higher education. [5,6,10]

Based on the results of this research, the following scientific-methodological and practical recommendations were developed:

For the Ministry of Preschool and School Education and Other Authorized Bodies:

Establish a "Center for Implementing Artificial Intelligence Technologies in Education" under the Ministry of Preschool and School Education of the Republic of Uzbekistan and the Ministry of Higher Education, Science, and Innovation. The functions of this center should include:

- Developing a strategy for applying AI technologies in education;
- Certifying AI-based educational platforms and assessment systems;
- Training and enhancing the qualifications of pedagogical staff in AI technologies;
- Studying foreign experience and adapting it to local conditions.

Develop and approve regulatory documents for implementing AI-based assessment systems in educational institutions, including:

- A standard on "Technical and Methodological Requirements for AI-Based Assessment Systems";
- Regulations on "The Procedure for Officially Recognizing Results of AI-Based Assessment";
- Guidelines on "Data Security and Personal Data Protection in AI-Based Assessment."

Implement pilot projects for introducing AI-based assessment systems. When selecting projects, it is recommended to consider the following criteria:

- The level of technical equipment of the educational institution;
- The qualifications of the pedagogical staff and their readiness for innovation;
- The infrastructural capabilities of the region;
- Coverage of various educational levels (primary, secondary, secondary special, higher).

For Heads of Educational Institutions and Pedagogical Staff:

Implement the AI-based assessment system in stages. Recommended stages:

- Preparation stage (3-6 months): Establishing technical infrastructure, training pedagogical staff, conducting explanatory work with students and parents.
- Trial stage (6-12 months): Testing the system in a limited number of classes or subjects, analyzing results, identifying and correcting shortcomings.
- Wide-scale implementation stage (1-2 years): Applying the system in all classes and subjects, monitoring, and continuous improvement.

Establish a system for improving the qualifications of pedagogical staff in AI technologies. Recommended training modules:

- "Fundamentals of Artificial Intelligence and its Potential in Education";
- "Working with AI-Based Assessment Systems";
- "Analyzing and Interpreting Assessment Results";
- "Forming Personalized Learning Paths with AI Assistance."

Enhance students' digital literacy and develop skills for working with AI systems, including:

- Introducing special modules on AI fundamentals in informatics and information technology lessons;
- Organizing an elective subject on "Artificial Intelligence and Future Professions";
- Establishing clubs and groups focused on AI technologies;
- Conducting AI project competitions among students.

For Researchers and Developers:

Continuously improve and update the developed AI models. Recommended directions:

- Testing more complex computer vision architectures (Vision Transformer, Swin Transformer) for mosaic assessment models;
- Fine-tuning multilingual models (mBERT, XLM-RoBERTa) for crossword assessment models and creating specific versions for the Uzbek language;

- Developing semantic models specialized for assessing open-ended responses in Uzbek for test assessment models.

Expand and enrich datasets for training AI models. Recommended measures:

- Creating a large database of mosaic, crossword, and test tasks in the Uzbek language;
- Developing high-quality, labeled datasets with the involvement of expert pedagogues and methodologists;
- Making datasets available as open-source for researcher use;
- Artificially expanding datasets through synthetic data generation.

Improve the system architecture and enrich it with new capabilities. Recommended directions:

- Creating the possibility to add new types of tasks (simulations, projects, portfolios) to the system;
- Integrating the system with Virtual and Augmented Reality (VR/AR) technologies;
- Developing a mobile version of the system and adapting it to various mobile devices;
- Ensuring the reliability and security of assessment results using blockchain technology in the system.

For Students and Their Parents:

Recommendations for students on effectively using the AI-based assessment system:

- Analyzing recorded actions and time parameters within the system to improve personal work efficiency;
- Using error analysis to identify and fill knowledge gaps;
- Following personalized recommendations and planning the learning process;
- Tracking personal progress through the system and analyzing it regularly.

Recommendations for parents on using the AI-based assessment system:

- Regularly obtaining information about the child's mastery level and progress through the system;
- Discussing recommendations provided by the system with the child and assisting in their implementation;
- Monitoring the child's interaction with the system and providing support when necessary;
- Comparing the child's results not with other students, but with their own previous results, and encouraging personal development.

Conduct long-term studies on the impact of AI-based assessment systems on the quality of education. Research directions:

- The impact of AI-based assessment on student motivation and academic achievement;

- The effectiveness of the system across different age groups and subject areas;

- The correlation between AI-based assessment results and traditional assessment results;

- The impact of the system on students' metacognitive skills and independent learning abilities.

Develop international cooperation and study advanced practices. Recommended measures:

- Participating in international projects in the field of AI and education;

- Conducting joint research with foreign scholars and research centers;

- Presenting research results at international conferences and seminars;

- Publishing articles in international scientific journals.

Promising Directions for Further System Improvement:

Create a Multimodal Assessment System. In the future, it is advisable to enrich the system with multimodal models capable of analyzing not only visual and textual data but also audio and video information. This will allow for assessing students' oral responses, presentations, and practical skills.

Incorporate Capabilities for Assessing Emotional Intelligence. Analyzing the student's emotional state (stress, interest, fatigue) to enable assessment and tailored support that considers their psycho-emotional condition.

Utilize Neurointerface Technologies. Exploring future possibilities for analyzing student brain activity via neurointerfaces to determine attention levels, cognitive load, and difficulty levels.

Expand Predictive Analytics Capabilities. Creating models based on the large volume of data collected by the system to predict students' future academic performance, learning difficulties, and their causes.

Conclusion. As a result of the conducted research, a system for the comprehensive AI-based assessment of mosaic, crossword, and test tasks was developed and scientifically substantiated. The research results fully confirmed the hypotheses initially proposed: the complex use of artificial intelligence technologies (computer vision, large language models, agent-based systems, semantic analysis) increases the objectivity and reliability of assessment, allows for the assessment of not only the final result but also the student's thought process, creates the possibility of providing personalized recommendations adapted to students'

individual characteristics, increases the predictive capability of assessment results, and saves teachers' time and effort.

The developed system aligns with the goals and objectives of the state policy on modernizing, digitizing, and implementing AI technologies in Uzbekistan's education system and serves to improve the quality of education, account for students' individual characteristics, and develop modern competencies.

References:

1. Resolution of the President of the Republic of Uzbekistan, No. PQ-4623, February 27, 2020, "On Measures for the Further Development of Pedagogical Education."

2. Rasch, G. (1960/1980). “Probabilistic models for some intelligence and attainment tests”. Copenhagen: Danish Institute for Educational Research. (Expanded edition, Chicago: University of Chicago Press, 1980). https://en.wikipedia.org/?curid=2161615

3. The Rasch model is considered a "significant milestone" in psychometrics and played a key role in the transition from Classical Test Theory to Item Response Theory (IRT). https://rd.springer.com/chapter/10.1007/978-3-030-43469-4_2

4. Lord, F. M., & Novick, M. R. (1968/2008). “Statistical theories of mental test scores”. Reading, MA: Addison-Wesley. (Reprinted by Information Age Publishing, Charlotte, NC, 2008). https://www.syndetics.com/index.aspx/?client=tenna&isbn=9781607529613/summary.html&type=rn12

5. Abdumominova, S. S. qizi, & Akbaraliyeva, A. (2025). Internetdagi testlar: Yangi muammolari va istiqbollari [Online Tests: New Challenges and Perspectives]. “Наука и технология в современном мире”, 4(3), 55–59. DOI: https://doi.org/10.5281/zenodo.14848055

6. Ovxunov, I.A., & Mamadaliyeva, K.X. (2026). Oliy ta'limda baholash jarayonini avtomatlashtirishda sun'iy intellekt texnologiyalarining qo‘llanishi [Application of Artificial Intelligence Technologies in Automating the Assessment Process in Higher Education]. “Zenodo”. https://doi.org/10.5281/zenodo.18352542

7. Pardayev, S. M. (2023). Possibilities of using digital platforms for assessing individual educational achievements of students. “Journal of Advanced Scientific Research” (ISSN: 0976-9595), 3(10).

8. Pardayev, S. (2023). Use of pedagogical, information and communication technologies in effective organization of the educational process. “Journal of Foreign Languages and Linguistics”, 6(2).

9. Mavlyanovna, K. S., Ogli, O. A. A., Ulashovich, T. F., Karshiboevich, S. S., & Mamasharipovich, P. S. (2021). An Integrated Approach To Information Technology Training In Non-Specialised Education. “Int. J. of Aquatic Science”, 12(3), 813-819.

10. Pardayev Sherzod Mamasharipovich. (2023). Modern approaches to the assessment of individual learning results of students. “Qokon University Newsletter”, 9(9), 193–196. https://doi.org/10.54613/ku.v9i9.870.