AI-Driven Personalized Learning in Higher Education: Balancing Adaptivity and Academic Rigor

Andrey Dragunov[1], Anna Tomskova[2]

[1]Inha University in Tashkent, Tashkent, Uzbekistan
Email: *a.dragunov[at]inha.uz*
https://orcid.org/0000-0001-6137-4602

[2]New Uzbekistan University, Tashkent, Uzbekistan
Email: *a.tomskova[at]newuu.uz*
https://orcid.org/0009-0006-9174-7127

Abstract: ***Artificial intelligence has significantly expanded personalized learning opportunities in higher education through adaptive instruction, automated feedback, and individualized pacing. At the same time, the widespread use of generative AI has challenged the validity of traditional assessment practices. This paper examines both the educational benefits and the pedagogical risks of AI-driven personalization. It discusses major adaptation mechanisms used in higher education and analyzes risks associated with over-adaptivity, including dependency, illusion of competence, reduced originality, and weakened assessment authenticity. The study also presents a practice-based case study from mathematics education in which graded homework became unreliable as evidence of independent student understanding because of increasing AI-assisted solution generation. In response, assessment practices were redesigned to emphasize non-graded formative homework, in-class board work, and controlled written examinations. Based on this experience, the paper proposes a hybrid assessment framework that balances AI-supported learning with AI-resistant evaluation. The central argument is that reliable assessment in AI-rich educational environments requires triangulation across multiple evaluation contexts rather than dependence on a single assessment format. The proposed framework offers a practical conceptual model for preserving academic rigor while retaining the educational advantages of AI-supported personalization.***

Keywords: artificial intelligence, personalized learning, higher education, assessment, academic rigor, generative AI, mathematics education, adaptive learning, formative assessment, assessment redesign, educational technology, AI-assisted learning

1. Introduction

Artificial intelligence has moved personalized learning from a niche educational idea to a mainstream institutional priority. Adaptive systems can now recommend materials, adjust task difficulty, generate hints, and provide feedback at a scale that was previously impossible (Zawacki-Richter et al., 2019; Popenici & Kerr, 2017). In higher education, these capacities are attractive because they promise better support for diverse student backgrounds, uneven preparation, and different learning paces. Recent international discussions on AI in education have therefore focused not only on innovation, but also on governance, equity, academic integrity, and the preservation of academic standards (UNESCO, 2025; Lodge et al., 2024). AI-driven personalized learning typically relies on continuous analysis of learner data and educational analytics derived from student interaction patterns and assessment outcomes (Zawacki-Richter et al., 2019). Platforms monitor assessment results, interaction patterns, time spent on tasks, and other indicators to estimate current proficiency and recommend the next step. In principle, this helps weaker students receive additional scaffolding while allowing stronger students to move toward more advanced tasks. Such flexibility may improve student engagement and reduce learning frustration, particularly in large and academically heterogeneous courses (Popenici & Kerr, 2017), where instructors cannot continuously individualize instruction for every student.

The same technologies, however, create a serious pedagogical dilemma. Once generative AI becomes widely available, students can use it not only as a tutor, but also as a producer of polished answers. As a result, educational environments face a new tension: the tools that make learning more adaptive may also weaken the validity of traditional assessment. Homework, take-home tasks, and routine written exercises become increasingly difficult to interpret as evidence of independent understanding in the presence of generative AI systems capable of producing polished academic responses (Kofinas et al., 2025; Gruenhagen & Parker, 2024).

This paper addresses that tension directly. It has three main contributions. First, it provides a structured discussion of the educational value of AI-driven personalization in higher education. Second, it identifies major risks associated with excessive adaptivity and AI dependence, including the possibility that assessment becomes detached from actual student competence. Third, it develops a hybrid assessment framework, grounded in a real teaching case, for balancing AI-supported learning with AI-resistant evaluation.

2. Methodology

This study uses a practice-based exploratory case study design situated within reflective higher education research. The paper combines conceptual analysis of AI-driven personalized learning with observations derived from instructional practice in undergraduate mathematics education. The purpose of the study is not to establish causal statistical relationships, but to examine how widespread access to generative AI affects the reliability of assessment in university-level courses and to propose a practical assessment framework responsive to these challenges.

The empirical component of the study was conducted during the teaching of undergraduate Discrete Mathematics courses at New Uzbekistan University (Tashkent, Uzbekistan). The observations were collected across one academic semester from several course sections consisting primarily of first-year undergraduate students in STEM-related programs. Approximately 80–120 students participated across the observed cohorts. The instructional context included regular homework assignments, in-class board problem solving, and proctored written examinations.

The study employed qualitative classroom observation and reflective instructional analysis. Particular attention was given to differences between AI-assisted homework performance, public in-class problem solving, and controlled written examination results. Instructor observations were documented throughout the semester using grading records, classroom notes, and comparative evaluation of student performance across assessment formats. The analysis focused on recurring behavioral and performance patterns rather than formal experimental measurement.

The statement that "approximately 10%" of students performed substantially better on the written examination than during board assessment was derived from comparison of internal course assessment records. Specifically, students whose board assessment performance suggested partial mastery but whose controlled written examination results demonstrated substantially higher competence were identified through instructor review of grading outcomes. The figure is therefore approximate and descriptive rather than the result of formal statistical testing.

The research should be understood as exploratory and practice-oriented rather than fully quantitative. Its primary contribution lies in pedagogical interpretation and framework development grounded in authentic classroom experience.

All observations were based on normal instructional activities and anonymized course performance data. No personally identifiable student information was used. The study did not involve intervention beyond standard educational practice. Institutional ethical approval was not formally required under local policy for reflective pedagogical analysis of anonymized classroom observations; however, the study was conducted in accordance with general principles of academic confidentiality and responsible educational research.

3. AI Adaptation Mechanisms in Higher Education

AI systems adapt learning through a combination of data analysis, probabilistic modeling, and feedback generation. At a basic level, they infer what a student likely knows, where the student is struggling, and which intervention may be helpful next. The underlying techniques vary across platforms, but they commonly include supervised learning, reinforcement-based updating, natural language processing, and rule-based recommendation systems. In higher education, these tools are frequently integrated into existing learning management systems and adaptive educational platforms (Zawacki-Richter et al., 2019).

From a pedagogical perspective, the most important feature of such systems is responsiveness. Instead of presenting identical material to all students, adaptive platforms can sequence tasks according to demonstrated performance. A student who repeatedly makes the same algebraic error may receive additional examples, simpler intermediate steps, or targeted explanations; a student who answers confidently and correctly may be directed toward extension tasks or more complex applications. This dynamic adjustment is one of the main reasons AI has been embraced as a mechanism for scalable personalization in higher education (Popenici & Kerr, 2017).

AI tools also provide rapid formative feedback and support AI-assisted formative assessment practices (Furze et al., 2024). In language learning, they can analyze pronunciation or grammar patterns; in STEM courses, they can detect recurring procedural errors; in writing-intensive settings, they can help students revise structure and clarity. When used appropriately, these systems reduce the waiting time between attempt and response, which is especially valuable in large university classes. They also make it easier for instructors to identify common misconceptions across cohorts and to refine course design accordingly.

However, adaptation mechanisms are not pedagogically neutral. The way a system defines progress, difficulty, or mastery shapes the learning trajectory that students experience. For that reason, AI adaptation mechanisms should be understood not only as technical tools, but also as pedagogical choices that encode assumptions about what counts as understanding, when support should be given, and how challenge should be introduced.

4. Educational Benefits of Personalized AI Support

The educational value of AI personalization lies in its potential to combine scale with responsiveness (Selwyn, 2019). Universities frequently teach students with highly uneven levels of prior knowledge, language preparation, and study habits. A single instructional pathway rarely serves all students equally well. Personalized AI support can reduce this mismatch by allowing students to revisit foundational ideas, receive extra practice, and progress at a pace more appropriate to their readiness (Popenici & Kerr, 2017).

In this sense, AI can strengthen formative learning. Students who hesitate to ask questions in class may still engage with interactive explanations privately (Popenici & Kerr, 2017; Selwyn, 2019). Learners who need repetition can obtain it without waiting for office hours, while those who are ready for more advanced material can move beyond routine tasks. When these tools are used as supplements rather than replacements for instruction, they can improve continuity of study between classes and increase the amount of deliberate practice students perform.

A further advantage is analytic visibility. Aggregate

Paper ID: SR26503141339 DOI: https://dx.doi.org/10.21275/SR26503141339

educational analytics can reveal patterns that are difficult to detect from instructor intuition alone, including recurring misconceptions, disengagement trends, and ineffective scaffolding strategies (Zawacki-Richter et al., 2019). Used carefully, such information can help instructors redesign tasks, adjust pacing, and support students more effectively. For these reasons, AI-driven personalization can be pedagogically valuable even when it does not replace traditional teaching methods.

5. Risks of Over-Adaptivity and AI Dependence

The same features that make AI support attractive can become pedagogically harmful when they are not balanced by rigorous academic expectations. One major risk is the illusion of competence, in which students overestimate their understanding because AI systems continuously scaffold problem solving and provide polished explanations on demand (Lodge et al., 2024). If students receive continuous hints, simplified pathways, or polished explanations on demand, they may feel that they understand a topic more deeply than they actually do. This becomes especially visible when students encounter unfamiliar problems without assistance and discover that they cannot independently organize a solution.

A second risk is skill atrophy, where repeated reliance on AI systems may weaken independent reasoning and sustained problem-solving ability (Selwyn, 2019). Higher education does not only transmit correct answers; it develops habits of reasoning, persistence, and self-correction. When AI routinely performs the most difficult parts of the task, students may begin to outsource exactly those processes that education is supposed to strengthen. In mathematics and related fields, this can mean reduced fluency in constructing arguments, checking steps, or sustaining concentration through a multi-stage problem.

There are also broader ethical and institutional concerns related to transparency, unequal access to AI technologies, governance, and academic integrity (UNESCO, 2025; Luo, 2024). AI systems may reproduce bias, privilege students with better access to digital tools, or encourage opaque forms of dependence. In assessment contexts, the main challenge concerns authenticity and the evidentiary reliability of submitted student work in AI-rich environments (Kofinas et al., 2025; Gruenhagen & Parker, 2024). This is particularly serious in university courses, where grades are expected to represent independent knowledge, not merely access to technological support.

These risks do not imply that AI should be excluded from education. Rather, they show that personalization without carefully designed boundaries can weaken both fairness and rigor. The central challenge is therefore not whether to use AI, but how to structure learning and assessment so that support does not erase evidence of real understanding.

6. Exploratory Case Study: Assessment Authenticity in Undergraduate Mathematics

The assessment problem becomes especially clear in mathematically structured courses. In Discrete Mathematics, homework traditionally serves two roles: it gives students opportunities for practice and it provides instructors with evidence of individual progress. Under conditions of widespread AI use, these roles begin to separate. Students may still benefit from AI-assisted homework as practice, but the resulting written submissions increasingly lose value as assessment evidence. Solutions often become highly polished, stylistically homogeneous, and difficult to distinguish in terms of authentic student understanding.

Recent studies have similarly reported that generative AI complicates the interpretation of take-home assessment and increases pressure to redesign authentic evaluation practices (Kizilcec et al., 2024; Gruenhagen & Parker, 2024).

Within the instructional context described in the methodology section, a revised model of assessment was implemented at New Uzbekistan University. Homework was retained, but only as a formative component and not as a graded one. Summative classroom evaluation was shifted to in-class work with previously unseen problems. Students were asked to solve these tasks at the board, in real time and under instructor observation. This format made AI assistance impossible during the act of evaluation and allowed instructors to see not only the final answer, but also the student's reasoning process, procedural fluency, and degree of conceptual control.

This change strengthened the authenticity of assessment. Board work gave instructors access to independent student thinking in a way that written homework no longer could. At the same time, it also exposed a limitation that is pedagogically important. Many students performed below their actual knowledge level when working publicly at the board. Mathematical problem solving requires concentration, internal organization, and time to think through several steps. Under classroom pressure, some students became anxious, forgot material they had learned, or were unable to demonstrate abilities that were visible in less exposed settings.

A recent proctored examination provided evidence of this gap. Review of internal assessment records suggested that approximately 10% of students performed substantially better on the written exam than they had during board assessment. In some cases, students whose board performance suggested only partial mastery later demonstrated near-complete or complete command under exam conditions. Because board evaluation was conducted by experienced instructors and the written examination was sufficiently controlled to prevent cheating, both measures can be treated as reliable within their own contexts. The discrepancy therefore does not indicate failure of either format; rather, it shows that each captures a different aspect of student performance.

This case highlights a central trade-off of AI-era assessment design. Evaluation methods that are highly resistant to AI misuse may underestimate student ability because of stress, time pressure, or public exposure. Conversely, formats that better reflect stable performance may not fully reveal independence if they are too vulnerable to external assistance. Preserving rigor therefore requires not a single perfect instrument, but a combination of assessment contexts.

7. A Hybrid Assessment Framework

Building on the case above, this paper proposes a hybrid assessment framework for AI-rich learning environments. The framework is designed to balance three objectives: resistance to AI misuse, fair representation of student knowledge, and support for ongoing learning. It rests on the premise that no single assessment format is sufficient once AI becomes a routine part of students' academic lives.

The need for multimodal and authenticity-oriented assessment frameworks has become increasingly prominent in recent discussions of AI-supported higher education assessment (Furze et al., 2024; Luo, 2024).

The first component is AI-supported homework used in a formative, non-graded way. In this component, students may use AI tools to obtain explanations, practice procedures, compare approaches, and identify gaps in understanding. The educational value of homework is preserved, but its role changes: it becomes a space for learning rather than a source of direct grades. This acknowledges the reality of AI use without pretending that written homework still functions as a transparent measure of independent competence.

The second component is in-class board assessment (Luo, 2024). Here students solve previously unseen problems under direct instructor supervision. This format has high authenticity because it reveals independent reasoning, procedural control, and the ability to work without technological support. It is particularly useful for identifying whether a student can initiate and structure a solution in real time. Its limitation is that performance may be affected by pressure and public visibility.

The third component is the controlled written exam. Unlike board work, it gives students more time, more privacy, and a more stable environment for sustained reasoning, while still excluding AI assistance. Written exams therefore complement board assessment by capturing forms of competence that emerge more clearly when students are not required to perform publicly and immediately. Together, these two summative components provide a more balanced picture than either would alone.

The key principle of the framework is triangulation. Student achievement should be interpreted across multiple assessment contexts rather than inferred from a single source. Homework reflects engagement and supported learning; board work reflects independence and authenticity; written exams reflect stability of performance under controlled conditions. Used together, these measures reduce the bias that would arise if any one of them were treated as sufficient on its own.

In conceptual terms, the framework can be visualized as a triangular model whose three vertices are formative AI-supported homework, authenticity-oriented board assessment, and stability-oriented written examination. Balance is achieved not by maximizing one vertex, but by calibrating the relationship among all three.

8. Practical Implications for Universities

The proposed framework suggests several practical implications for universities. First, institutions should reconsider the assumption that all completed work can still be interpreted as evidence of independent student knowledge. In AI-rich environments, some forms of work are better treated as learning processes rather than grading instruments. This does not diminish their value; it changes the function they serve.

Second, instructors need assessment designs that distinguish between support and evaluation. If AI is allowed in practice settings, then summative formats must be intentionally designed to capture independent reasoning. The appropriate balance will vary across disciplines, but the underlying principle remains the same: learning can be AI-supported, whereas evaluation must contain clearly protected spaces of independence.

Third, institutions should train faculty not only in the operational use of AI tools, but also in assessment redesign appropriate for AI-rich educational environments (UNESCO, 2025). The pedagogical challenge is no longer simply how to use AI in teaching; it is how to preserve rigor, fairness, interpretability, and academic integrity while generative AI becomes normalized within higher education (Lodge et al., 2024).

9. Conclusion

Artificial intelligence has expanded the possibilities of personalized learning in higher education by improving access to adaptive instruction, feedback, and individualized academic support. However, the rapid normalization of generative AI has also weakened the reliability of traditional assessment formats as indicators of independent student competence. This paper argues that preserving academic rigor in AI-rich educational environments requires a transition from single-format evaluation toward integrated assessment models that combine multiple forms of evidence.

The presented case study demonstrates that no individual assessment method fully captures both authenticity and fair representation of student knowledge. In-class board work, controlled written examinations, and formative AI-supported homework each reveal different dimensions of student learning. The proposed hybrid framework therefore emphasizes triangulation across assessment contexts rather than dependence on a single grading instrument.

Although the framework requires further empirical validation across disciplines and institutional settings, it offers a practical and implementable approach for balancing AI-supported personalization with rigorous evaluation standards in higher education.

Table 1: Functional roles of the proposed hybrid assessment framework

Component	Primary purpose	Main strength	Potential limitation
AI-supported homework	Formative learning and exploration	Encourages practice, feedback, and guided revision	Low validity as direct evidence of independent performance
In-class board assessment	Authenticity-oriented evaluation	Reveals independent reasoning in real time	Sensitive to stress, time pressure, and public exposure
Controlled written exam	Stability-oriented evaluation	Captures sustained reasoning under secure conditions	Occurs less frequently and may not show immediate spontaneity

References

[1] Furze, L., Perkins, M., Roe, J., & MacVaugh, J. (2024). The AI Assessment Scale (AIAS) in action: A pilot implementation of GenAI-supported assessment. *arXiv*. https://arxiv.org/abs/2403.14692

[2] Gruenhagen, J. H., & Parker, R. (2024). The rapid rise of generative AI and its implications for higher education assessment: A scoping review. *International Journal of Educational Technology in Higher Education, 21*(40). https://doi.org/10.1186/s41239-024-00453-6

[3] Kizilcec, R. F., Huber, E., Papanastasiou, E. C., Cramb, A., Makridis, C. A., Smolansky, A., Zeivots, S., & Raduescu, C. (2024). Perceived impact of generative AI on assessments: Comparing educator and student perspectives in Australia, Cyprus, and the United States. *Computers and Education: Artificial Intelligence, 7*, 100269. https://doi.org/10.1016/j.caeai.2024.100269

[4] Kofinas, A. K., Tsay, C. H.-H., & Pike, D. (2025). The impact of generative AI on academic integrity of authentic assessments within a higher education context. *British Journal of Educational Technology*. https://doi.org/10.1111/bjet.13585

[5] Lodge, J. M., Howard, S., Bearman, M., & Dawson, P. (2024). The evolving risk to academic integrity posed by generative artificial intelligence. Australian Tertiary Education Quality and Standards Agency (TEQSA).

[6] Luo, J. (2024). A critical review of GenAI policies in higher education assessment: A problem representation approach. *Assessment & Evaluation in Higher Education*. https://doi.org/10.1080/02602938.2024.2309963

[7] Popenici, S. A. D., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning, 12*(1), 22. https://doi.org/10.1186/s41039-017-0062-8

[8] Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education*. Polity Press.

[9] UNESCO. (2025). *Artificial intelligence and digital education: Policy guidance for higher education*. UNESCO Publishing.

[10] Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education, 16*(39). https://doi.org/10.1186/s41239-019-0171-0