Task Design and Assessment Strategies for AI-Influenced Education

Vira Liubchenko[1,*,†], Nataliia Komleva[1,†] and Svitlana Zinovatna[1,†]

[1] *Odesa Polytechnic National University, 1 Shevchenko av., Odesa, 65044, Ukraine*

Abstract

The rapid integration of generative AI (GenAI) into higher education presents both opportunities and risks for authentic learning. While AI can enhance efficiency and personalization, it also threatens academic integrity by enabling superficial task completion and diminishing cognitive engagement. This paper proposes an information technology framework designed to minimize the adverse influence of GenAI while preserving its educational benefits. The methodology introduces temporally shifted assignments, content-break micro-assessment cycles, and differentiated analysis of AI's impact on student outputs. The approach is scalable, programmatically implementable, and compatible with existing learning management systems, making it more sustainable than labor-intensive safeguards such as oral examinations. Empirical validation across five software engineering courses demonstrated improvements in task authenticity, student comprehension, and critical thinking, while reducing reliance on AI-generated solutions. The results confirm that structured task design and iterative teacher–student interactions foster deeper engagement and enhance the reliability of learning outcomes. The study underscores the need to move beyond purely ethical guidance toward technological safeguards integrated into instructional design. Future research will focus on adaptive platforms capable of dynamically embedding this methodology across diverse curricula and monitoring the depth of student engagement with AI systems.

Keywords

artificial intelligence, information technology, educational task design, assessment strategies, learning management systems

1. Introduction

The rapid development of artificial intelligence (AI) and its widespread integration into various domains of human activity have led to profound transformations in the educational environment. Generative AI (GenAI) technologies, particularly large language models (LLMs), have become accessible to a wide range of students and educators, offering new opportunities to support the learning process. However, alongside the potential for personalization, rapid access to knowledge, and automation of routine tasks, the use of AI has introduced several significant challenges to the education sector.

Contemporary students actively utilize AI tools to complete academic tasks, including text generation, problem-solving, query formulation, and response structuring. At the same time, there is a growing concern regarding the declining depth of understanding of educational material, which is manifested in limited abilities for analysis, synthesis, and reflection. These trends raise concerns about achieving the stated learning outcomes and preserving a meaningful educational process. On the other hand, educators often lack appropriate tools or methodological frameworks for effectively integrating AI into their teaching practices.

The learning process can be viewed as an information process, where each stage is associated with the handling of data packets. The teacher collects information on the course domain and

ICST-2025: Information Control Systems & Technologies, September 24-26, 2025, Odesa, Ukraine

* Corresponding author.
† These authors contributed equally.
lvv@op.edu.ua (V. Liubchenko); komleva@op.edu.ua (N. Komleva); zinovatnaya.svetlana@op.edu.ua (S. Zinovatna)
0000-0002-4611-7832 (V. Liubchenko); 0000-0001-9627-8530 (N. Komleva); 0000-0002-9190-6486 (S. Zinovatna)

processes it, resulting in a structured presentation of theoretical content and a set of assignments for the practical component. The student processes the theoretical content to form knowledge and then applies that knowledge to complete practical tasks, thereby developing skills. Finally, the student uses this knowledge to perform summative assessments, the results of which provide data for the teacher to evaluate the level of achievement of the intended learning outcomes.

The uncontrolled use of AI has led to significant disruptions in this informational process. At the initial stage, teachers can utilize AI tools to enhance the efficiency of course development. In the subsequent stages, however, students may delegate both practical and summative tasks to AI systems, which process the theoretical content on their behalf. As a result, the outputs being assessed are no longer authentic indicators of the student's learning outcomes.

A strategy of rejecting AI entirely is neither viable nor aligned with labor market demands, as the intelligent use of AI enhances task efficiency and productivity. Therefore, it is necessary to restructure or reconfigure the educational process in a way that utilizes AI to provide added value rather than posing a threat to academic integrity.

In this context, there is a pressing need for a systematic analysis of AI's impact on learning outcomes, the identification of critical risks, and the development of approaches for modifying educational tasks and curricula.

2. AI challenges in education

The active integration of AI technologies into educational processes is accompanied not only by increased efficiency in specific learning components but also by the emergence of multiple challenges that affect the quality of knowledge acquisition, pedagogical interaction, and the achievement of learning outcomes.

In [1], a bibliometric analysis of AI applications in education concludes that while pedagogical aspects were considered before 2020, more recent publications have increasingly focused "on the technical aspects of implementing AI rather than on pedagogical models that could underlie its use in education."

A student survey analyzed in [2] revealed that "95.6% of respondents use AI in their academic activities, underscoring the deep integration of this technology into modern education. Virtual assistants are the most commonly used AI tools (88.2%), providing support for information retrieval, task management, and real-time feedback". AI adoption strategies and use cases vary across regions. For instance, [3] outlines practical approaches to integrating AI in Algerian higher education, including generating instructional prompts, designing multi-step assignments, and utilizing GenAI tools to enhance student engagement. In [4], the challenges of AI in Islamic religious education for senior secondary school students in Indonesia are discussed.

According to [5], 90% of students reported that, despite AI's potential in education, their teachers did not encourage its use as a learning aid. Rather than leveraging tools like ChatGPT for group or class projects, students mostly used it for individual assignments. Furthermore, students stated that they were not instructed on how to use ChatGPT safely and effectively.

As GenAI technologies continue to evolve, ongoing research and adaptable instructional strategies are crucial for maximizing benefits while mitigating potential drawbacks [6].

Numerous studies have examined both the potential benefits and the associated risks of implementing AI in education (e.g., [7–9]). Below, we highlight several of the documented advantages of AI within the educational context.

Personalized and adaptive learning systems. Intelligent tutoring systems (ITS), as described in [10], are capable of answering students' questions in real time, delivering immediate feedback and support, and helping students solve complex learning problems. As [11] reports, students using ITS demonstrated higher learning quality than those using traditional methods, although some findings predate the public availability of advanced AI tools. Notably, [12] emphasizes that "despite the vast amount of information AI can analyze regarding student achievements and personal preferences, nothing can replace the human educator's ability to observe emotional cues and build meaningful emotional connections with students".

Automated and semi-automated assessment systems have been developed to enhance student learning outcomes by providing timely and constructive feedback. As demonstrated in [13], LLMs can effectively support educators in conducting comprehensive and methodologically validated assessments of student responses when fine-tuned for specific domains. In this context, [14] highlights that teaching quality improvement strongly depends on resource-oriented approaches, which laid the groundwork for later AI-driven assessment systems. Similarly, [15] emphasizes a paradigm shift in knowledge evaluation, where automated exam systems ensure objectivity and transparency, while also raising questions of reliability and trust in AI-based assessment. Furthermore, [16] provides a systematic review of trends in AI-driven education, classifying modern applications such as adaptive learning, automated grading, and ethical challenges, thus offering a structured overview of the field.

Additionally, AI-based systems can monitor the educational process and detect potential issues at an early stage. By analyzing academic performance and behavioral data, such systems enable the timely identification of students who require additional support, thereby assisting teachers and administrators in delivering targeted interventions [17].

However, the integration of AI into education also introduces a range of potential risks and challenges for stakeholders. One frequently noted concern is the decline in critical thinking and cognitive skills, as students may increasingly rely on AI for quick answers rather than engaging in independent learning and analytical reasoning [18]. As highlighted in [19], a central challenge lies in striking the right balance between leveraging the advantages of AI and fostering the development of fundamental cognitive abilities; notably, 83% of surveyed respondents expressed the belief that overreliance on AI could significantly impair their capacity for independent thought. Additional concerns relate to the reliability of assessment, as automated grading and AI-generated content complicate plagiarism detection and make it more challenging to verify student authorship [16]. The issue of student dependency on AI is also critical, since excessive reliance may hinder the cultivation of independent problem-solving and reasoning skills [16]. Ethical dimensions further compound these risks: while [13] underscores that AI's educational potential cannot be separated from the need for responsible use, [20] observes mounting concern over the misuse of systems such as ChatGPT in educational contexts, though it remains unclear whether such practices affect broader ethical attitudes within the sector. Finally, questions of equity must also be considered, as AI systems may unintentionally reinforce inequalities in access to educational opportunities, particularly if training data fail to represent diverse learner populations [16] adequately.

As stated in [8], educational institutions must urgently "update academic integrity policies and plagiarism guidelines". Faculty members need to be educated on AI tools, while students should be made aware of the responsible use of AI and its potential implications for academic integrity.

Suppose a teacher fails to detect unethical AI use, resulting in inflated grades. In that case, the lack of material comprehension goes unnoticed, which undermines the effectiveness of the educational process and jeopardizes the true purpose of teaching and learning [9].

The initial enthusiasm surrounding the potential of AI in education has gradually subsided, and since 2024, an increasing number of studies have shifted their focus toward risks, challenges, and possible mitigation strategies. However, much of this literature concludes with broad calls for change while offering few concrete pathways for implementation. The present study addresses this gap by formalizing the problem and advancing structured approaches to its resolution.

3. Problem statement

The learning process can be conceptualized as an informational cycle that involves two primary information-processing agents: the teacher and the student. This cycle unfolds through a series of sequential stages: (1) the teacher collects information regarding the current state of knowledge within the subject domain as well as the level of students' preparedness; (2) based on this information, the teacher develops instructional materials and designs tasks for both practice and assessment; (3) the student engages with the materials and completes practice tasks to acquire and strengthen practical skills; (4) the student undertakes assessment tasks to demonstrate the achievement of learning outcomes; (5) the teacher evaluates the completed tasks and updates the information about students' preparedness; and (6) the process recommences from the first stage.

The integration of information technologies has considerably enhanced the efficiency of this cycle. Nonetheless, the advent of GenAI has introduced critical challenges. The uncontrolled use of GenAI has resulted in situations where students delegate the execution of training and assessment tasks to automated systems (Figure 1). Such practices undermine the authenticity of the learning process, as teachers are no longer evaluating students' genuine knowledge and competencies but rather AI-generated outputs.

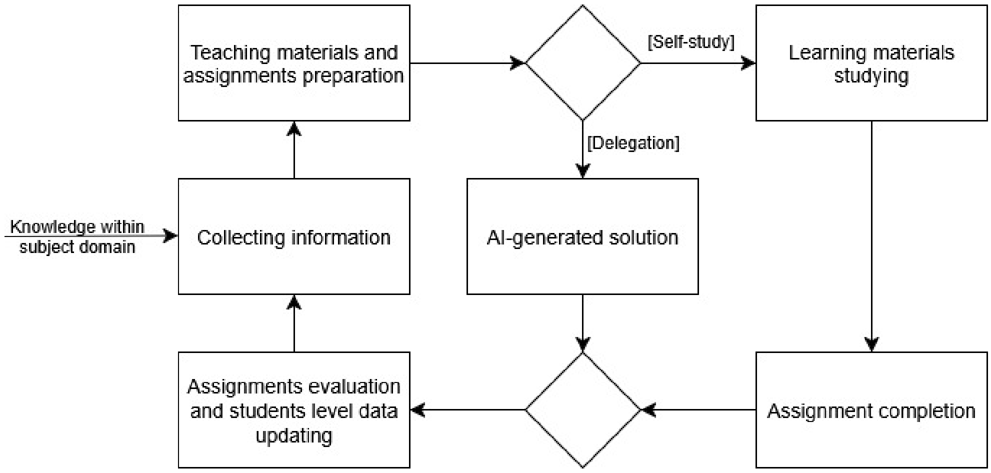

Figure 1: Schema of the learning process

To mitigate this problem, current approaches emphasize strengthening direct student–teacher interaction. These approaches include oral examinations, individual questioning during the submission of practical assignments, and the systematic review of students' intermediate project outcomes by teachers. While these measures may enhance authenticity, they present substantial scalability challenges. For instance, applying them in large cohorts of approximately 150 students entails disproportionate time demands on teachers, thereby limiting their practicality.

Consequently, there is a need to develop novel methods for constructing packages of assessment tasks so that their completion cannot be outsourced to AI tools without the active involvement of the student. Such methods would contribute to safeguarding the validity and reliability of learning outcomes in the context of widespread access to generative AI technologies.

4. Modeling the educational process with minimization of generative AI influence

Traditional instructional models are often based on a linear (waterfall) approach in which the teacher assigns a task, the student completes it by the deadline, and the teacher evaluates the result. An educational program (EP) within an academic institution consists of a set of educational components (EC) and corresponding learning outcomes (LO), as defined in the curriculum:

$$EP = \langle EC, LO \rangle.$$

Each educational component $ec_i \in EC$ supports a subset of learning outcomes $LO_k \subseteq LO$, denoted as $ec_i \to LO_k$. Conversely, each learning outcome $lo_j \in LO$ is supported by a subset of educational components $EC_j \subseteq EC$, i.e., $lo_j \to EC_j$.

Additionally, for each learning outcome lo_j addressed in an educational component ec_i, there may exist a set of sub-learning outcomes SLO_i not explicitly specified in the EP, but which are necessary for achieving lo_j in ec_i: $SLO_i \Rightarrow lo_j$ (Figure 2).

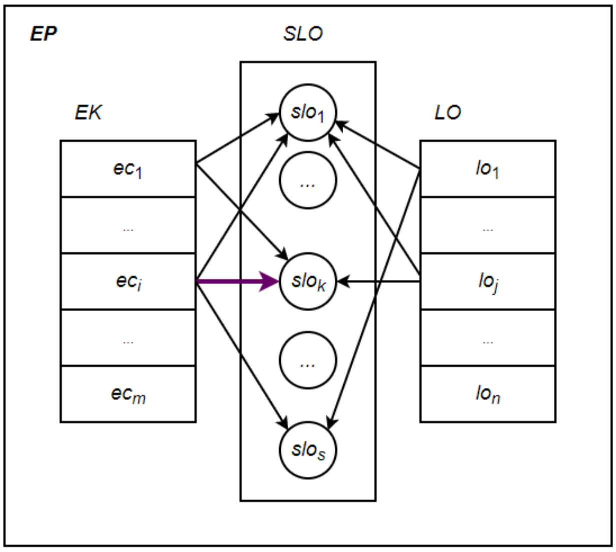

Figure 2: Relationship between educational components and learning outcomes within an educational program

This approach becomes increasingly less effective in digital environments where students can consult GenAI systems at any stage to receive a complete solution without engaging in deep cognitive processes.

To mitigate the influence of GenAI, we propose a unified methodology that combines two complementary mechanisms: temporal separation of tasks (time breaks) and structural separation of task content (content breaks). Together, these mechanisms are intended to increase task resistance to automation while ensuring a more authentic demonstration of students' knowledge.

To verify the achievement of SLO_i within a specific ec_i, the teacher should provide appropriate tasks ct_l. Each task ct_l passes through three life-cycle stages: task creation by the teacher, solution generation by the student, and evaluation by the teacher. Currently, each of these stages can be performed by AI tools. Therefore, it is essential to design tasks that minimize AI influence on the demonstration of student-acquired knowledge slo_k during the study of ec_i. This can be formally expressed as:

$$ct_l{}' = \arg\min_{iai_i} slo\_demo(ct_i, iai_i), \tag{1}$$

where $slo\_demo()$ is the function representing the demonstrable outcomes of a task for assessment, ct_i is the task, and iai_i denotes AI influence on the result.

In the presence of AI tools, not only the task content but also the interaction model between participants of the educational process must be transformed. As recommended in [21], tasks should be constructed to elicit knowledge directly obtained by the student, rather than solutions generated by AI.

The teacher, therefore, faces the challenge of designing tasks that align with these principles. Drawing on the source content of educational components, ec_i, used in forming control tasks for learning, ct_l, the following classification can be proposed (Figure 3):

a) tasks based exclusively on new content not previously employed in assignments for the given EC;
b) tasks drawing on content that has already been used within the same EC;
c) tasks integrating content from assignments associated with other ECs.

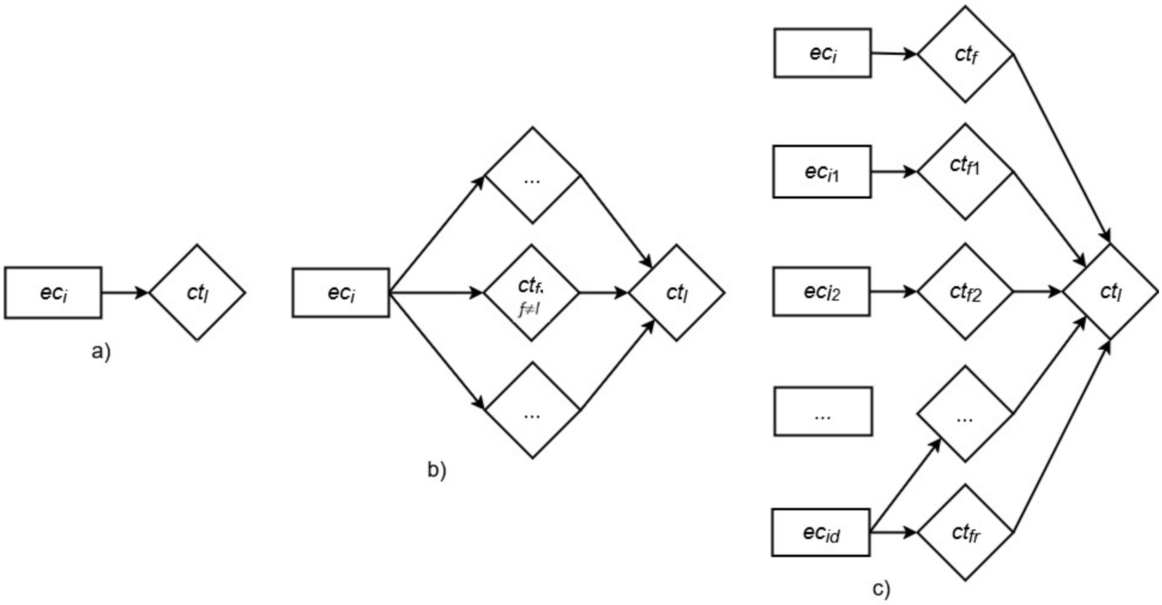

Figure 3: Task classification according to temporal separation: a) the task draws exclusively on new content; b) the task utilizes content that has already been covered within this EC; c) the task integrates content from assignments associated with other ECs.

The inclusion of references to past assignments within current tasks introduces a concept known as task time break (*brt*). Accordingly:

- class (a) is considered a 0-degree time break (brt_0),
- class (b) is a 1st-degree time break (brt_1),
- class (c) is a 2nd-degree time break (brt_2).

In modern learning management systems such as Moodle, tasks can be annotated with custom metadata fields that describe their origin and temporal relationship to prior assignments [22]. Using the `Custom fields API` or dedicated plugins (e.g., `Custom fields for activity modules`), teachers may define fields such as `EC_id`, `Content_source`, and `Related_assignments`. Based on these values, the system can automatically classify the task into one of the three categories (brt_0, brt_1, brt_2).

This classification can be encoded algorithmically, as shown below:

```
def classify_task(content_source, EC_current, EC_history):
    if content_source not in EC_history:
        return "brt0"
    elif content_source in EC_history[EC_current]:
        return "brt1"
    else:
        return "brt2"
```

The resulting metadata not only provides transparency but also supports integration with content break assessment. Teachers can link each task to predefined outcomes or competencies within Moodle, enabling rubrics such as *slo_demo*() to incorporate the degree of task time break directly. This creates an operational bridge between the conceptual model and real-world teaching environments.

Introducing time breaks in tasks helps reduce AI influence on student outputs by:

- requiring students to apply prior knowledge from earlier tasks,
- increasing the complexity of AI queries needed to solve multi-context problems,
- activating non-local knowledge that may be less accessible to AI.

Complementary to time breaks, content breaks involve the systematic fragmentation of tasks into smaller, meaningful components accompanied by intermediate reviews, oral justifications, and iterative refinement. In contrast to waterfall-style assignments that concentrate solely on the final product, the content-break approach emphasizes cyclical interaction between teachers and students. Each cycle produces partial results, followed by discussion and feedback, thereby reducing the potential for AI-driven automation and fostering authentic student engagement.

The overall score within this model can be calculated as:

$$Score_{total} = W_{sprint} * Score_{sprint} + W_{review} * Score_{review} , \qquad (2)$$

where $Score_{total}$ denotes the overall (final) score; W_{sprint} and W_{review} are the weighting coefficients assigned to sprint and review components, respectively (e.g., 0.6 and 0.4); $Score_{sprint}$ represents the average score for tasks completed within a sprint, adjusted by the corresponding task time break level

$$Score_{sprint} = \frac{1}{m}\sum_{j=1}^{m}\left(BaseScore_j * \left(1 + \alpha * brt_{level\,j}\right)\right) ; \qquad (3)$$

m indicates the number of tasks within a sprint; $BaseScore_j$ is the base score of the j-th task; α is the scaling factor that regulates the influence of the time break (typically 0.1–0.2); $brt_{level\,j} \in \{0,1,2\}$ designates the task time break level of the j-th task; $Score_{review}$ refers to the average score for demonstrated learning outcomes, such as oral or written reviews

$$Score_{reviw} = \frac{1}{n}\sum_{i=1}^{n} slo\_demo(ctl_i) ; \qquad (4)$$

n specifies the number of SLOs assessed during the review phase; $slo\_demo(ctl_i)$ is the score assigned for demonstrating the i-th learning outcome during review; ctl_i represents the specific content or learning task segment associated with the assessed outcome.

Content-break task design facilitates cyclical interaction between teachers and students, enabling systematic monitoring, feedback, and the adaptive refinement of subtasks. A complex assignment is decomposed into a sequence of subtasks, $p(ct_i)$, aligned with the planned strategy for its overall solution. These subtasks are presented sequentially, encouraging students to follow the necessary progression of steps toward a coherent outcome. While students may employ AI tools to address individual subtasks, active participation is required to ensure consistency across the entire solution. Each subtask is evaluated either by the teacher or by an AI assistant, and the outcomes of this evaluation provide the basis for targeted recommendations and, when necessary, corrections to subsequent subtasks.

In this context, it is essential to rethink the types of learning tasks used to form and assess learning outcomes. Waterfall-style tasks emphasize the final product, making them highly susceptible to AI automation. In contrast, tasks requiring step-by-step development, justification of decisions, discussion of interim results, and progressive refinement significantly complicate full automation and encourage genuine student engagement.

5. Educational process information technology under the influence of generative AI

In response to the challenges posed by the uncontrolled use of GenAI in the educational process, an information technology has been proposed to enhance the authenticity of student task performance and support the achievement of learning outcomes. The proposed technology models the learning process as a dynamic system of interactions among educational components, learning outcomes, supporting learning outcomes, and task types, all of which must be adapted to the digital context.

Within this framework, a sequence of steps is defined, from the formalization of the educational program to the monitoring of learning achievements. Each stage is supported by control mechanisms that aim to reduce the risk of formal or automated task completion without sufficient cognitive involvement from the student. Key mechanisms include the use of temporally shifted tasks (time break, *brt*), content-break assessment, and differentiated analysis of AI influence on the completion of individual tasks.

The structure of the proposed information technology is presented in Table 1, which outlines the implementation stages, their functional objectives, the stakeholders involved in the educational process, and the corresponding verification mechanisms.

The proposed technology enables the transformation of the educational process to meet modern digital challenges. Through the implementation of cyclic control mechanisms, task time shifts, and dynamic performance assessment, it significantly reduces the risk of superficial or automated learning facilitated by GenAI tools.

Table 1
Stages of the proposed methodology

Stage	Description of Actions	Control/Verification Mechanism
1. Formalization of the *EP*	Define the educational program as a set of *EC* and *LO*. Establish relationships between them and identify *SLO*.	Methodological analysis, validation of *EC*–*LO* mappings
2. Design of Tasks with Time Break (*brt*)	Create ct_l tasks using content with varying levels of time break to increase difficulty for automated solutions.	Inclusion of cross-topic links in tasks; coherence validation
3. Development of ct_l with Partial Control	Decompose each ct_l into parts $p_d(ct_i)$ for cyclic evaluation according to content-break principles (weekly/bi-weekly checkpoints).	Evaluation of intermediate task components with a focus on student progress
4. AI Influence Assessment	Assess the impact of GenAI on the outcomes of each ct_l using the function $slo\_demo(ct_i, iai_i)$.	Comparison of problem-solving logic, analysis of queries, and behavior patterns
5. Review and Feedback	Perform regular review of task parts (ct_i), provide feedback, and adapt subsequent tasks accordingly.	Real-time teacher feedback; student explanation of the reasoning process
6. Analytics and Quality Control	Aggregate statistics on ct_l completion, AI usage, LO achievement, and the effectiveness of *brt* and content-break mechanisms.	Data collection and dynamic monitoring via an analytics system

A key advantage of this technology lies in its adaptability to various levels of cognitive complexity and its ability to ensure transparent interaction among participants in the learning process.

6. Case studies

The proposed methodology for designing practical and assessment tasks was validated within the educational programs of the Information Technology field. Here, we provide a detailed description of the pilot implementations.

Case 1. Course "Fundamentals of Programming" (1st semester of study)

The study involved 142 students, divided equally between a control and an experimental group. The control group completed traditional algorithmic tasks such as sorting, searching, and array manipulation. In contrast, the experimental group undertook a modified assignment organized in three stages: manual tracing of algorithms, coding without reliance on built-in functions, and an oral defense of the proposed solution. The educational interventions thus combined paper-based algorithm tracing, restrictions on the use of pre-defined functions, and verbal justification of implementation logic. Student performance was evaluated according to three criteria: their understanding of algorithmic principles, ranging from no explanation to correct reasoning fully; the correctness of program implementation, assessed on syntax, error-free execution, and compliance with task requirements; and the quality of oral explanation, varying from superficial description to clear, well-reasoned argumentation. The findings demonstrated a marked improvement in the experimental group, where 72% of students were able to explain algorithms effectively compared to only 38% in the control group. Statistical analysis confirmed a strong effect ($\chi^2 = 26.5$, $p < 0.001$, Cramer's $V = 0.52$).

Case 2. Course "Prompt Engineering" (3rd semester of study)

The study involved 98 students, divided equally between control and experimental groups. The control group was instructed to design a single prompt for generating a step-by-step guide to cloud infrastructure setup. In contrast, the experimental group was required to produce at least three distinct prompts, evaluate the resulting AI-generated responses against selected criteria, and compose a written reflection on the observed differences. The educational intervention, therefore, combined the development of multiple prompts, the systematic assessment of AI outputs, and reflective analysis of the outcomes. Student performance was evaluated in terms of the completeness of AI responses, their correctness and technical accuracy, the practical applicability of the generated solutions, and the depth of student reflection, with each criterion scored on a five-point scale. The results revealed a marked improvement in the experimental group, where 63% of students demonstrated the ability to differentiate relevant from flawed responses, compared with only 30% in the control group. Statistical analysis confirmed a medium effect ($\chi^2 = 18.9$, $p < 0.001$, Cramer's $V = 0.36$).

Case 3. Term Paper in Software Systems Engineering (6th semester of study)

The study involved 138 students, divided equally between control and experimental groups. The control group completed a term project using a traditional approach, working independently for one to two months. In contrast, the experimental group followed a content-break model incorporating sprint-based reviews. In this model, the project was divided into stages covering requirements, architecture, implementation, testing, and documentation. Progress was reviewed in bi-weekly sprint sessions with the teacher, and the feedback obtained was systematically integrated into subsequent stages. Student performance was assessed according to the quality of requirements, the soundness and originality of the architecture, the correctness and completeness of code and testing, and the structure and professionalism of the documentation. The results demonstrated that only 45% of students in the control group displayed a clear understanding of the software development life cycle, compared with more than 70% in the experimental group. Project quality in the experimental group increased by 12% on the grading scale, while reliance on template-based solutions declined threefold. Statistical analysis confirmed a strong effect ($\chi^2 = 24.7$, $p < 0.001$, Cramer's $V = 0.51$).

Case 4. Course "Databases" (4th semester of study)

The study was conducted over five academic years (2021–2025) and included 419 students in total. The control group comprised students from 2021 to 2024 ($n = 312$), who completed assignments without additional constraints, while the experimental group consisted of students in 2025 ($n = 107$), who were required to provide an oral defense of their work. The intervention thus introduced a mandatory oral explanation during project defense while maintaining the same assignment structure without temporal separation. Student performance was evaluated according to the correctness of ER model design, the accuracy and optimization of SQL queries, and the quality of oral justification, each assessed on a five-point scale. In 2024, the average score for the second assignment reached its highest level at 87% of the maximum. After the introduction of oral defense in 2025, the average adjusted to 79%, indicating more differentiated and authentic performance. These results suggest that student outcomes became less dependent on uncontrolled AI use and more reflective of genuine knowledge. Statistical analysis confirmed a medium effect ($\chi^2 = 13.7$, $p < 0.001$, Cramer's $V = 0.34$).

Case 5. Course "Data Modeling and Visualization" (5th semester of study)

The study involved 84 students, divided evenly between a control group and an experimental group. The control group completed both assignments in the standard format. In contrast, the experimental group worked under a scheme that introduced temporal separation and context-aware task design, with the second assignment explicitly linked to datasets previously analyzed by the students. The evaluation focused on four dimensions: the quality of research plan design, assessed on a scale from unstructured approaches to fully logical sequencing; the integration of prior results, ranging from no connection to complete alignment with earlier work; the depth of data interpretation, spanning from general remarks to detailed analysis and reasoned conclusions; and the authenticity of responses, evaluated from generic AI-generated outputs to unique, contextually grounded work. The results showed that the experimental group produced more structured and context-sensitive assignments, whereas the control group often relied on generic, AI-like outputs. Statistical analysis confirmed a medium effect ($\chi^2 = 21.3$, $p < 0.001$, Cramer's V = 0.42).

A comparative overview of performance improvements across all five case studies is shown in Figure 4.

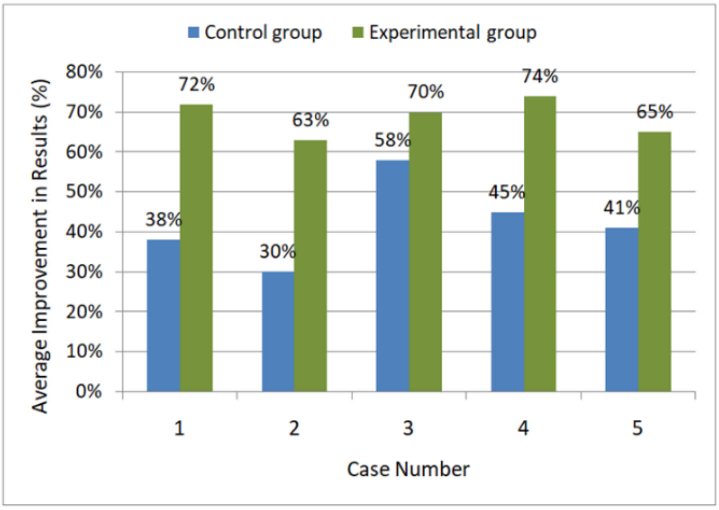

Figure 4: Comparative bar chart of performance across case studies

Table 2 provides a summary of all key statistical indicators across the five case studies.

Table 2
Statistical results of educational interventions across case studies

Case	χ^2	p-value	Cramer's V	Interpretation
Case 1: Fundamentals of Programming	26.5	<0.001	0.52	Strong effect
Case 2: Prompt Engineering	18.9	<0.001	0.36	Medium effect
Case 3: Course Project in Software Systems Engineering	24.7	<0.001	0.51	Strong effect
Case 4: Databases	13.7	<0.001	0.34	Medium effect
Case 5: Data Modeling and Visualization	21.3	<0.001	0.42	Medium effect

The most pronounced improvements were observed in Case 3, where regular sprint reviews enhanced overall project quality and substantially reduced reliance on template-based solutions. Case 1 also demonstrated significant gains, with redesigned assignments markedly improving students' understanding of algorithmic principles.

Moderate yet positive effects were identified in Case 2 and Case 4, where students developed greater capacity for critical evaluation and oral articulation of task solutions. Case 5 further confirmed that context-linked assignments effectively distinguished engaged students from those who relied primarily on AI-generated outputs.

The forest plot (Figure 5) summarizes the effect sizes (Cramer's V) across all five case studies, with 95% confidence intervals depicted for each. Dashed reference lines at 0.3 and 0.5 indicate thresholds for medium and strong effects, respectively, facilitating a clear comparison of the relative impact of the educational interventions. This visualization highlights which cases yielded moderate improvements and which achieved a strong effect.

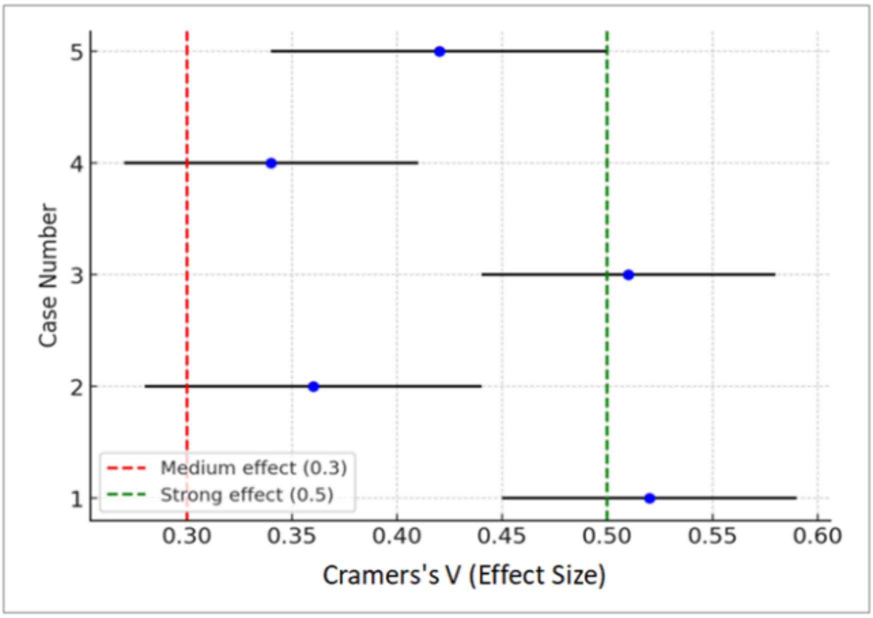

Figure 5: Forest plot of Cramer's *V* across case studies

To ensure objective assessment, it is essential to examine the distribution of results, as this reflects not only the average level of knowledge but also the variability in task performance across different groups.

Figure 6 presents the outcomes of five database tasks for both the control and experimental groups. The visualization conveys not only the mean values but also the distribution of results: the boxes represent the interquartile range (25–75%), the red lines indicate the median, and the whiskers denote the minimum and maximum values, excluding outliers.

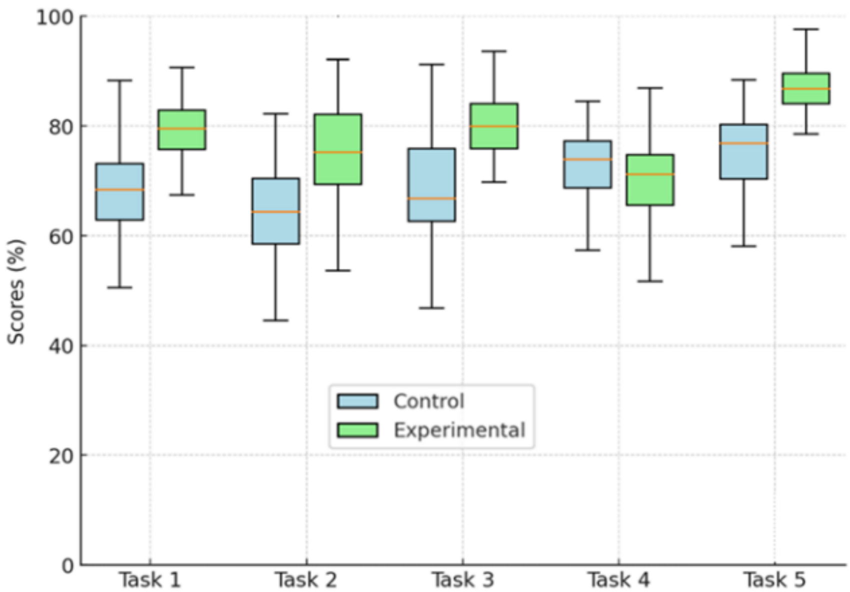

Figure 6: Examples of variation in student scores

In most cases, the experimental group achieved higher scores, albeit with greater dispersion, reflecting individual differences in mastering the material following the introduction of new requirements. For instance, in Task 4, the average score in the experimental group is slightly lower, attributable to the increased challenge of orally justifying ER models and SQL queries. This suggests that the revised approach more accurately captures students' true understanding, rather than merely measuring mechanical completion or reliance on pre-generated solutions or AI assistance.

Ultimately, it is essential to emphasize that the implementation and evaluation of the proposed approach were conducted primarily within Information Technologies programs. Accordingly, its applicability to other academic disciplines, particularly those with differing pedagogical objectives or assessment formats, such as the humanities, visual arts, or medical education, has yet to be established.

7. Conclusion

The rapid integration of GenAI tools into higher education necessitates a fundamental rethinking of traditional approaches to learning, assessment, and instructional design. While these technologies offer unprecedented opportunities, their widespread availability also poses considerable risks, including the acquisition of superficial knowledge, diminished cognitive engagement, and distortions in the evaluation of student performance.

A review of existing publications on the incorporation of GenAI tools into the learning process indicates that scholarly attention has primarily focused on technologies for designing instructional strategies, particularly steps 1–2 and 5 outlined in Section 3. Ethical training for students is generally emphasized at steps 3–4, under the assumption that informed learners will restrict their use of GenAI to approved purposes. However, this assumption has proven untenable. Leading universities increasingly report that such an ethics-based approach is insufficient, as evidenced by the reintroduction of oral examinations and other safeguard measures.

The methodology proposed in this study does not dismiss the importance of AI ethics education. Instead, it advocates a shift from purely moral–ethical safeguards to technological safeguards embedded within the learning process itself. By designing assignments that contain semantic or temporal discontinuities, the methodology generates negative feedback when students rely exclusively on GenAI, thereby discouraging disengagement and reinforcing the necessity of active human participation.

A notable strength of this approach is its scalability. The methodology can be programmatically implemented and seamlessly integrated into existing learning management systems. Unlike oral examinations, it does not require disproportionate amounts of teacher time, thus preserving a valuable academic resource while maintaining instructional rigor.

The effectiveness of the methodology was validated across multiple academic disciplines. Empirical evidence demonstrates improvements in the quality of student outputs, enhanced awareness of task execution logic, and a reduction in formulaic or AI-generated responses. Furthermore, the introduction of micro-assessment cycles, supported by regular student–teacher interactions, was shown to foster critical thinking and reflective capacities.

Future research should focus on the development of adaptive digital platforms capable of embedding this methodology dynamically across diverse educational programs. Particular attention ought to be directed toward automated tools for monitoring student interactions with AI systems and assessing the depth of their cognitive engagement throughout the learning process.

Declaration on generative AI

During the preparation of this work, the authors used Grammarly to check grammar and spelling. After using this service, the authors reviewed and edited the content as needed and took full responsibility for the publication's content.

References

[1] V. Gonzalez Calatayud, P. Prendes, R. Roig-Vila, Artificial Intelligence for Student Assessment: A Systematic Review, Applied Sciences 11.12 (2021) 5467. doi: 10.3390/app11125467.

[2] A. M. Vieriu, G. Petrea, The Impact of Artificial Intelligence (AI) on Students' Academic Development, Education Sciences 15(3) (2025) 343. doi: 10.3390/educsci15030343.

[3] S. Hamane, S. Khalki, AI in education: transforming the teaching profession and unlocking future opportunities in Algeria, Journal for Educators, Teachers and Trainers 15 (2024) 69–78. doi:10.47750/jett.2024.15.04.007

[4] S. Syahrizal, F. Yasmi, M. Thomson AI-Enhanced Teaching Materials for Education: A Shift Towards Digitalization, International Journal of Religion 5 (2024) 203–217. doi: 10.61707/j6sa1w36.

[5] S. Chakraborty, Generative AI in Modern Education Society, arXiv (2024). doi: 10.48550/arXiv.2412.08666.

[6] S. Krause, B. H. Panchal, N. Ubhe, The evolution of learning: Assessing the transformative impact of generative AI on higher education, arXiv (2024). doi: 10.48550/ARXIV.2404.10551.

[7] M. Özer, Potential Benefits and Risks of Artificial Intelligence in Education, Bartın University Journal of Faculty of Education 13(2) (2024) 232–244. doi: 10.14686/1416087.

[8] C. K. Lo, What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature, Education Sciences 13(4) (2023) 410. doi: 0.3390/educsci13040410.

[9] S. Grassini, Shaping the Future of Education: Exploring the Potential and Consequences of AI and ChatGPT in Educational Settings, Education Sciences 13(7) (2023) 692. doi: 10.3390/educsci13070692.

[10] Q. Xu, Action Research Plan: The impact of the use of Artificial Intelligence in Education on the Cognitive Abilities of University Students, Springer Science and Business Media LLC, Discover Education 3 (2024) 224. doi: 10.21203/rs.3.rs-4981281/v1.

[11] L. Chen, P. Chen, Z. Lin, Artificial Intelligence in Education: A Review, IEEE Access 8 (2020) 75264–75278. doi: 10.1109/ACCESS.2020.2988510.

[12] S. Saylam, N. Duman, Y. Yildirim, K. Satsevich, Empowering education with AI: Addressing ethical concerns, London Journal of Social Sciences (2023) 39–48. doi: 10.31039/ljss.2023.6.103.

[13] E. Kasneci, ChatGPT for good? On opportunities and challenges of large language models for education, Learning and Individual Differences 103 (2023) 102274. doi: 10.1016/j.lindif.2023.102274.

[14] N. Komleva, V. Liubchenko, S. Zinovatna, Improvement of teaching quality in the view of a resource-based approach, CEUR-WS 2740 (2020) 262–277.

[15] A. Sharma, M. Khan, S. Patel, AI-based examination system: A paradigm shift in education sector, International Journal of Emerging Trends in Engineering Research 10(12) (2022) 29–36. doi: 10.30534/ijeter/2022/0410122022.

[16] M. Ratul, T. Rahman, S. Sultana, The role of AI in transforming education: A systematic review of trends, Education and Information Technologies 30(1) (2025) 55–79. doi: 10.1007/s10639-025-11987.

[17] M. O. Dei, The impact of AI on the adaptation of educational materials and teaching methods to the needs of each student, LatIA 3 (2025) 124. doi: 10.62486/latia2025124.

[18] S. Zhang, X. Zhao, T. Zhou, J. H. Kim, Do you have AI dependency? The roles of academic self-efficacy, academic stress, and performance expectations on problematic AI usage behavior, International Journal of Educational Technology in Higher Education 21 (2024) 34. doi: 10.1186/s41239-024-00467-0.

[19] K. Szmyd, E. Mitera, The Impact of Artificial Intelligence on the Development of Critical Thinking Skills in Students, European Research Studies Journal XXVII (2024) 1022–1039. doi: 10.35808/ersj/3876.

[20] N. Silva, I. Alvarez, Use and abuse of AI – ethical perspectives in the educational sector, The leading role of smart ethics in the digital world, Jorge Pelegrín-Borondo, Mario Arias-Oliva, Kiyoshi Murata, Mar Souto Romero (Eds.) (2024) 233–242.
[21] G. Siemens, F. Marmolejo-Ramos, F. Gabriel, K. Medeiros, R. Marrone, S. Joksimovic, M. de Laat, Human and artificial cognition, Computers and Education: Artificial Intelligence 3 (2023) 100107. doi: 10.1016/j.caeai.2022.100107.
[22] Moodle, Custom fields API, Moodle Developer Documentation (2025). URL: https://moodledev.io/docs/5.1/apis/core/customfields (accessed 17 August 2025).