Redefining Assessment strategies in Generative Artificial (GAI) Era

Reham Salhab[1,*]

Abstract: A thorough review of traditional student assessment strategies is necessary in light of the expanding use of generative artificial intelligence (GAI) in education. To ascertain their efficacy and applicability in the rapidly evolving educational environment in GAI era, this reevaluation is essential. This study aims to identify strategies for redefining assessment tools at a higher education institution. The criteria used to measure student achievement must be reevaluated to ensure that they appropriately represent the complexity of contemporary education, since GAI tools facilitate differentiated instruction and individualized learning. A qualitative study is conducted with a case study of a group of college instructors who use GAI tools in assessing students learning. 45 college instructors were recruited from three different colleges to answer the questions of semi structured interviews. Three main themes emerged: redesigning assessment tools, tasks redesign, and focusing on learning. Recommendations include that college instructors need to develop GAI assessment literacy. Limitations include generalizability due to small number of participants.

Keywords: General Artificial Intelligence (GAI), assessment, learning, personalized learning, assessment strategies.

إعادة صياغة استراتيجيات التقويم التربوي في ظل الذكاء الاصطناعي التوليدي

رهام سلحب[1,*]

الملخص: الهدف: تهدف هذه الدراسة إلى تحديد استراتيجيات إعادة تعريف أدوات التقييم في مؤسسات التعليم العالي في ظل انتشار الذكاء الاصطناعي التوليدي(GAI) ، وضمان ملاءمة معايير قياس التحصيل الدراسي للتعليم المعاصر الذي يتسم بالتمايز والتخصيص. **المنهج:** أجريت دراسة نوعية باستخدام منهج دراسة الحالة، حيث تمت مقابلة 45 مدرساً جامعياً من ثلاث كليات مختلفة باستخدام مقابلات شبه منظمة. ركزت الأسئلة على تجاربهم في استخدام أدوات الذكاء الاصطناعي التوليدي لتقييم الطلاب. **أهم النتائج:** كشفت الدراسة عن ثلاثة محاور رئيسية: 1. إعادة تصميم أدوات التقييم لتتناسب مع عصر الذكاء الاصطناعي. 2. إعادة تصميم المهام التعليمية لتعكس المهارات العليا للتفكير. 3. التركيز على عملية التعلم بدلاً من النتائج فقط. **الاستنتاجات والتوصيات:** أظهرت النتائج ضرورة تطوير كفاءة المدرسين في تقييم الطلاب باستخدام الذكاء الاصطناعي التوليدي (GAI Assessment Literacy). كما أوصت الدراسة بتطوير معايير تقييم مرنة تعكس التعلم الفردي. من أهم محدوديات.

الكلمات المفتاحية: التعلم، التقويم التربوي، الذكاء الاصطناعي التوليدي، التعلم المشخصن، استراتيجيات التقويم.

1 Technology education department, Arts and educational Sciences College, Palestine Technical University-Kadoorie, Tulkarm, Palestine
* Corresponding author email: r.salhab@ptuk.edu.ps

1 قسم الإرشاد والتوجيه، كلية التربية، جامعة ابن خلدون، إسطنبول، تركيا.
* الباحث المراسل: r.salhab@ptuk.edu.ps

Introduction

Assessment procedures in higher education institutions (HEIs) around the world have been severely disturbed by the advent of generative artificial intelligence (GAI) tools.

Even though, it seems that generative artificial intelligence (GAI) is emerging as a competitive factor among higher education institutions that endeavor to accomplish the United Nations' Sustainable Development Goals (SDGs), particularly SDG 4 (quality education) and SDG 5 (gender equality) (Awadallah Alkouk & Khlaif, 2024; Nedungadi *et al*., 2024), it still needs re-evaluation of its potentials. For example, as the integration of generative artificial intelligence (GAI) is advancing in different educational contexts, which necessitates a rethinking and redesigning of traditional educational assessment methods that prompts educators to analyze their effectiveness in today's revolutionized learning environment (Rudolph *et al*., 2023; Sabet *et al*., 2024). This scrutiny is essential, as GAI tools enhance personalized learning experiences and facilitate differentiated instruction, compelling us to revisit the metrics and guidelines that define student performance in ways that authentically represent the complexities of contemporary education (Baidoo-Anu& Ansah, 2023).

It is essential to assess the substance, scope, and appropriateness of these guidelines because many of them were developed in a hurry and affect a broad spectrum of teachers and students (Sesay, 2024).

Research gap

Integrating GAI applications into education has been widely studied (Foung *et al*., 2024; Madland at al., 2024:). Hence, a gap remains in understanding GAI integration into assessment practices, especially in higher education. While existing research primarily focuses on AI tools for grading or feedback, limited studies has been investigated how AI can transform methodologies of the assessment process (Xia *et al*., 2024).

Moreover, Gen AI in education studies is concentrated rarely conducted Palestine (Khlaif *et al*, 2025). This study seeks to address these gaps by exploring what assessment strategies could be used to adapt to GAI to minimize the challenges being addressed when using GAI for. This study focuses on college instructors who uses GAI assessment tools in different specializations.

Significance of the study

Redefining and critically reevaluating assessment procedures is crucial, to ensuring that they meet the demands of contemporary education as educational environments change due to the incorporation of technology like generative artificial intelligence (GAI) (Khlaif *et al*., 2025). The complexity of student learning and engagement is sometimes overlooked by traditional models, which are frequently focused on standardized testing (Jin *et al*., 2025). This is especially true in AI-enhanced learning settings that cater to a variety of demands (Al Harrasi *et al*., 2024).

A comprehensive approach to reimagining assessments must account for the varied learning modalities enabled by technology while ensuring these assessments are equitable, valid, and accurately reflect students' true abilities (Crawford *et al*., 2023). By drawing on emerging trends and innovative practices, educators can design assessments that foster critical thinking, creativity, and problem-solving skills, moving beyond the reliance on rote memorization and information recall (Vetrivel *et al*., 2025).

A systematic review of the integration of AI with the Internet of Things highlights the pivotal role of intelligent data analysis in

influencing educational outcomes. It suggests that tailored assessment strategies.

This study will significantly help educators in improving learners' experiences and foster meaningful knowledge retention while using GAI in assessment (Marengo, 2024). Thoughtfully redesigned assessments can provide timely feedback and enable more personalized learning journeys (Khalif *et al*., 2024). This shift is critical for preparing students to adapt to an ever-evolving workforce while supporting their unique educational paths (Khlaif, 2024). Equipping students with the necessary skills and competencies to succeed in a complex world is essential. Therefore, the transformation of assessment practices is not only about leveraging modern technology's potential in education, but also about improving the overall quality and effectiveness of learning (Saihi, 2024). Moreover, this study will offer guidance for policy-makers on how best to leverage the opportunities to assess students learning while using GAI in higher educational institutions. Theoretically, this study will contribute to the literature by addressing the gap in using effective assessment strategies while using GAI tools in college settings. Meanwhile, existing literature focuses mainly on AI's tools usage in the teaching learning process, this study aims to redefine assessment strategies while using GAI tools.

Hence this study aims to identify and suggest assessment strategies while using GAI tools at a higher education institution.

Literature review

Generative AI in Education

Technological advancement in deep learning and natural language processing (NLP) have produced complex models that can generate quality text, graphics, and various types of content, among other tasks. By automating and simplifying procedures, these models—collectively referred to as Generative Artificial Intelligence (GAI), have the ability to completely revolutionize education by composing music, creating images or writing texts, in a way that, in principle, would be indistinguishable from a human creation. Models such as GPT-4 (Liu *et al*., 2023) or LLaMA 2 (Kabal *et al*., 2024), are representative examples of this type of technology, capable of generating coherent and contextual texts based on specific guidelines.

Formally, GAI can be defined as the creation of unpublished synthetic material, in any format, to assist with any activity, via generative modeling (García-Peñalvo & Vázquez-Ingelmo, 2023). In the face of the new scenarios that arise with GAI, there are shadows of misuse and superficial learning (Jawabreh & Itmazi, 2025), which are directly related to assessment activities. However, it is also true that GAI offers the possibility of optimizing and automating some of these assessment processes, freeing up time and resources (Ng *et al*., 2024). But, and more importantly, it proposes an innovation in assessment paradigms, so that they can be more personalized, adaptive and, ultimately, fairer (Mohammed & Rocke, 2024).

GAI provides promising solutions to achieve the SDG4 goals, centering on inclusiveness equitability learning environments, and quality education for all (Wang & Zang, 20241). For example, GAI is focusing on the ability to personalize learning environments that addresses diverse needs by minimizing dropout rates and enhancing learning outcomes which achieves SDG 4 goal (Nedungadi *et al*., 2024). Moreover, GAI offers diverse instructional materials in different languages, which supports less advantaged students and requires strategic investments for accessibility. Further, GAI also has the potential to offer access for education. GAI can offer access underserved regions by enabling distribution of instructional material to students

who cannot access the conventional educational institutions physically. Correspondingly, this inclusiveness offers equitable learning opportunities and lessens educational inequities (Kumar et el., 2023).

In the educational field, the influence of these technologies is palpable and constantly evolving. However, before delving into the innovations and changes driven by AI, it is essential to understand how assessment processes have typically worked, with special emphasis on pre-university educational levels.

Because GAI technologies, like ChatGPT, have been pre-trained on large datasets, they can take context into account to customize responses to users' unique demands, like enhancing learning experiences (Samala *et al.*, 2024). ChatGPT is a potent instrument for revolutionizing education is generative artificial intelligence (GAI). Unprecedented degrees of individualized learning have been made possible by GAI applications, including data augmentation, intelligent tutoring systems, content creation, personalized material, and automated grading tools.

Unprecedented degrees of individualized education have been made possible by GAI applications, including data augmentation, intelligent tutoring systems, automated grading software, content development, and tailored content (Saihi, 2024). By using these technologies, educators can tailor lessons to each student's needs, giving them the best chance to succeed (Saputra, 2024). For example, AI-driven methods present encouraging prospects for dynamic and interactive learning environments in spite of these obstacles. Education institutions may take use of AI's promise to empower students, enable individualized learning, and equip them for success in a world that is becoming more complicated by striking a balance between

technical innovation and critical oversight (McDonald *et al.*, 2025).

Furthermore, there are notable global changes in education that correspond with the growth of GAI (Ahmed *et al.*, 2024). The focus of recent revisions is on assisting students in acquiring the skills necessary to use their knowledge to solve issues and provide explanations in the real world.

Nevertheless, these technologies have significant drawbacks as well, such as ingrained societal biases in the training datasets (Mao *et al.*, 2024). When using GAI in educational settings, these constraints need to be carefully considered.

Moreover, the efficiency of GAI in meeting the varied requirements of students is called into doubt, though, as it may also homogenize educational experiences. Targeted interventions and improved engagement are made possible by AI's capacity to analyze vast amounts of data, yet dependence on such systems needs to be carefully considered. For example, conventional evaluations may still provide insights that AI tools might overlook (Pan *et al.*, 2024). Additionally, concerns about algorithmic bias, reduced human oversight, and the potential for fostering over-reliance on technology warrant careful evaluation (Boukherouaa *et al.*, 2021).

But there are also significant drawbacks to these tools, such as the presence of social biases in the datasets that were used to train them (Mao *et al.*, 2024). When using GAI in educational contexts, it is important to carefully consider these constraints.

Educational Assessment

In order to measure student comprehension, inform instructional decisions, and guarantee that learning objectives are being met, educational assessment is essential (Al-Ramamne *et al.*, 2017; Miller *et al.*, 2025). In order to complement instructional techniques

and enhance the learning process, meaningful feedback should be given.

Standardized tests, written exams, and quizzes are examples of traditional assessment techniques that are frequently used to gauge students' knowledge and proficiency. These methods, which concentrate on assessing learning outcomes at the end of a course or instructional time, usually fit within a summative assessment model. Although somewhat successful, traditional approaches have been heavily criticized throughout time.

These tests' strict, one-size-fits-all design is a significant flaw since it frequently ignores students' varied skill levels and learning preferences (Memarian & Doleck, 2024). Standardized assessments, for instance, could ignore Standardized examinations may neglect individual differences, including cultural backgrounds and distinct learning needs, and frequently offer minimal feedback, hindering pupils' ability to recognize and address their learning deficiencies (Koseda et al., 2025).

Furthermore, traditional assessments tend to prioritize lower-order thinking skills, such as memorization and recall, rather than promoting higher-order cognitive processes like critical thinking, creativity, and problem-solving (D'Agostino et al., 2023). The grading process for these assessments is another area of concern, as it can be time-intensive and susceptible to human error or bias, raising questions about fairness and reliability (Rezai, 2022). These limitations emphasize the need for more flexible, personalized, and scalable assessment strategies. Additionally, dissatisfaction with many of the traditionally used assessment methods was already evident with the lockdown during the COVID-19 pandemic (García-Peñalvo et al., 2020). When the arrival of the GAI puts back on the table the opportunity and need to rethink and innovate in assessment, with an approach closer to

learning-oriented assessment for empowerment (Khlaif, 2024). The nature of the defined assessment tasks must be taken into consideration, in which the existence of the GAI cannot be ignored (García-Peñalvo, 2024)

Generative AI and educational assessment

Standardized examinations may neglect individual differences, including cultural backgrounds and distinct learning styles. On the other hand, generative Artificial Intelligence (GAI) has significant potential to revolutionize education by leveraging its ability to consider context and deliver rapid (Cheng et al., 2024), meaningful evaluations of student learning outcomes by personalized and individualized learning (Arslan et al., 2024; Wu& Mo, 2025). However, current GAI tools face notable limitations, including the presence of social biases embedded in the datasets used to train them (Gökoğlu, 2024). This revolution coincides with a shift away from memorization-based education systems toward approaches that prioritize the application of knowledge and skills to address real-world problems and explain real-world phenomena (Kaldaras et al. 2024). Moreover, GAI offers a way to lower tests' costs both formative and summative use (Mao et al., 2024; Owoseni et al., 2024). A study conducted by Garcia-Panalove (2024) revealed the benefits of GAI in assessing learning processes that can be completed more quickly, allowing teachers to spend more time on instruction and personalized feedback. Personalized assessments that are created by GAI has the ability to analyze individual student performance and needs to generate assessments tailored to each profile. This not only provides a more individualized learning experience, but also helps to identify specific areas for improvement. Flores-Vivar & García-Peñalvo (2023), also investigated how Generative AI supports automation and efficiency of assessment by allowing teachers to automate

repetitive and laborious tasks. Also, the role of generative AI helped to make a transition from transitional, conventional, standardized testing procedures to more individualized tools with more comprehensive details of students' knowledge and abilities because these tools can be adapted to learners' needs (Khalif *et al.*,2024). Similarly, Vetrivel *et al.* (2024) explained the transformative impact of AI on assessment practices within education. By automating key elements of the assessment process, AI unlocks new possibilities for efficiency, customization, and scalability. AI-driven tools streamline grading, freeing educators to concentrate on offering personalized feedback and supporting student development. Moreover, AI's ability to process large datasets enables the identification of performance trends, facilitating targeted interventions and adaptive learning experiences designed to meet each learner's unique needs. Additionally, a study of Awadallah Alkouk and Khlaif (2024) explores how AI can enrich traditional assessment approaches and suggested that AI based assessment should foster critical thinking, creativity, and collaboration and emphasized on evaluating student's interaction instead of the final product. It appears from the literature review that some studies mentioned limitations of standardized examination. For example, Cheng *et al.* (2024) and Khalif *et al.* (2024) argue that standardized tests often fail to account for individual differences, such as cultural backgrounds and learning styles. These assessments prioritize uniformity over personalization, which can disadvantage students with diverse needs. On the other hand, some studies reported the advantages of GA as discussed by Arslan *et al.* (2024) and Wu & Mo (2025), GAI offers a more nuanced approach by leveraging context-aware evaluations and personalized learning pathways. GAI tools can analyze individual performance and generate

tailored assessments, providing a more equitable and individualized learning experience. Moreover, personalization and efficiency were reported by Garcia-Panalove (2024) and Flores-Vivar & García-Peñalvo (2023) who emphasize GAI's ability to automate repetitive tasks, such as grading, allowing educators to focus on instruction and personalized feedback. This automation not only reduces costs (Mao *et al.*, 2024; Owoseni *et al.*, 2024) but also enhances the efficiency of formative and summative assessments.

Also, adaptive learning was reported by Vetrivel *et al.* (2024) who highlight GAI's capacity to process large datasets and identify performance trends, enabling adaptive learning experiences and targeted interventions. This aligns with the shift toward application-based education systems, as noted by Kaldaras *et al.* (2024).

More specifically, GAI applications can be employed for formative assessments effectively by providing students with continuous and constructive feedback (Elmourabi *et al.*, 2024).

Nevertheless, GAI tools usage in assessment came with concerns and controversies. For example, ChatGPT is used to create test or writing assessment tasks, which decreases required resources that allows for the creation of more items. Additionally, ChatGPT frees-up faculty time to perform higher level assessment activities (Zucherman *et al.*, 2023). ChatGPT is also able to consistently produce items using a standard format while adhering to item writing guidelines, which can be very challenging for faculty teams. Some concerns related to exam validity and reliability is a concern because they are required qualities of any test. Moreover, Chatgpt creates incorrect items for assessment (Kolade *et al.*, 2024).

On the other hand, a study conducted by Floden (2025) who differentiated between

ChatGPT and humanized grading and reported that Chatgpt struggles to accurately score questions closely related to the content of course lectures while performs better on more general questions. This shows that GAI grading tools can generate slightly different grades than human graders. Meanwhile, GAI use for grading lead to ethical challenges as exams are less trusted to a machine whose decision-making criteria are not fully understood with potential bias in training data.

Meanwhile, using generative AI in assessment lacks providing quality and informative feedback. Effective measurement should evaluate high order understanding, specifically using knowledge to solve issues and provide explanations (Kaldaras *et al*., 2024).

Moreover, there is a major obstacle when utilizing GAI to score tests centered on knowledge application. Making sure that these algorithms measure the same qualities like knowledge and skills that human scorers are trained to assess. Similarly, when utilizing GAI to develop assessments, it is crucial to ensure validity of the assessment (Dowson *et al*., 2024).

Therefore, even if generative AI has enormous potential to transform the teaching learning process and improving institutional objectives, its integration will not solely require overcoming technical and financial barriers. Instead, it also needs addressing academic challenges that persist.

The theoretical framework of this study is reported by prior research (<u>**Lye & Lim, 2024**</u>; Khalif *et al*., 202**5**) who classified types of assessment to four categories (AAAE):The Against, Avoid, Adopt, and Explore (AAAE) framework. **Against (usage of non-AI assessments).** Avoid (negatives of GAI usage in assessment), Adopt (integration of GAI into assessments) and Explore (GAI partner in assessment).

It seems from the literature review that while GAI holds immense promise for transforming education, its implementation must be approached critically. Addressing issues of bias, ensuring alignment with modern educational goals, and fostering equitable access to GAI tools are essential for realizing its full potential. The studies collectively provide a robust foundation for understanding both the opportunities and challenges associated with GAI in education.

Currently, no certain guidelines have been created for validity of Gen AI-based assessments and their t results. The goal of this study is to identify strategies to improve the validity and quality of GAI-based assessments strategies. Hence this study is trying to address the following question:

What AI-assessment strategies could be used while using Generative artificial intelligence (GAI) at college level?

The following section will discuss the methods that were conducted to answer this question

Methodology

Research Design

A qualitative research approach was conducted to determine strategies used by college instructors in assessing students' performance while using generative AI. A qualitative approach helps to understand a phenomenon from the lived experience of a group of people (Mohajan, 2018). A case study was used as a research design, which offers an understanding of an experienced phenomenon in a certain context; this helps researchers to understand a real-life phenomenon (Yin, 2003). Case studies focus on one thing and do not seek to generalize from it (Thomas, 2011). Since this study targets a new experience and contextual conditions are highly relevant to the research

focus, case study is selected. Other methods like phenomenology and narrative will not to study the phenomena that is rarely presented.

Participants

The participants in the current study had to be college instructors in a higher education institution with three different branches, who have experience and familiarity with Gen AI tools, have integrated AI tools in teaching their courses, come from various fields and different campuses.

The study's participants included 45 college instructors from a technical university in Palestine, who had different backgrounds and various experiences of using Generative AI in assessing different subjects and disciplines as shown in table 1

Table (1): Participants characteristics.

Variable	Category	Frequency
Gender	Female	17
	Male	28
Degree	Associate Professor	17
	Assistant Professor	9
	Professor	9
	Lecturer	10
Technical skills	Excellent	26
	Very good	11
	Good	8
Major (Teaching faculty)	Arts and Social Sciences	10
	Applied Sciences	15
	Physical Education	4
	Business	9
	Information Technology	7

All of college instructors were required to use GAI in teaching and assessing their students.

Data were collected through semi-structured interviews conducted in the first semester of 2024/2025 via Zoom platform. The interviews lasted a median of 32 minutes, ranging from 20 to 42 minutes. An interview guide was developed by the author after scanning the literature related to GAI assessment startegies.

Recruitment

A maximum variation sampling was used to recruit the participants. This method allows researchers to select participants with different characteristics in order to comprehend a phenomenon by examining a range of different cases (Emmel, 2013). The criteria for selecting the participants were faculty members with different specialties.

Trustworthiness

Qualitative studies have multiple quality standards, including credibility, which corresponds to internal validity in quantitative studies, transferability, which corresponds to external validity or generalization, reliability, which corresponds to the concept of stability, and confirmability, which corresponds to the concept of objectivity (Seidman, 2013). The following is a presentation of the mechanism for achieving these standards:

For credibility participants were provided with a copy of the interview and allowing them to clarify any misinterpretations; which gave them the opportunity to add additional information related to the study. As for transferability, it was achieved by the researchers providing sufficient, comprehensive, detailed, and clear information about the topic under study and its tools and procedures, and providing a rich and accurate description of the study context that provides the reader with complete information about the context in which the study took place, which allows participants to replicate the study and thus facilitates the process of transferring the results to other similar situations. Transferability was also achieved through choosing the intended sample that contributed to achieving the goal of the study, and linking the results of the study to what was mentioned in the previous theoretical literature.

Reliability was achieved by the two researchers; reviewing the records of the interviews and focus groups that were transcribed, and reviewing all the reports and documents that were analyzed; to ensure the

clarity and transparency of the principle of reliability, in addition to the clarity of the study methodology, the design used, the method of implementation, the details of the sample, the method of its selection, and the documentation of the interviews and the two focus groups, the study became theoretically replicable by other researchers. Confirmation was achieved by fulfilling the previous criteria of using a clear description of the study methodology and design, clarifying the study procedures and the clarity of the role of the researchers, which demonstrated their lack of self-bias in the study procedures, interpreting the results, and reaching conclusions and results, and summarizing the content of each question asked during the semi structured interview. Data analysis is described, which confirmed that the results are not biased, but rather accurately depict the participants' responses.

Ethical considerations

This study received ethical approval from the university, and informed consent was obtained from all participants. Participants' identities were kept confidential, and their information was securely stored on a private computer accessible only to the researcher.

Data Analysis

This analysis focuses on identifying the pattern within the data, and inductive reasoning through repeated examination and comparison of data; To reduce the collected data to a set of topics or categories, and then generate, review, identify, and interpret knowledge (Levitt *et al*, 2018), where the objective analysis process of individual interviews, focus groups, reports, and documents went through the following stages:

Transcribing the recorded data, reading the written texts and rereading them more than once, and re-listening to the recordings; to take notes on the data while listening and while

reading again (from focus groups and individual interviews only).

Data coding (Coding): Start by writing initial codes for the data; this is done by carefully reading a passage or sentence and then summarizing it with one or two words.

Re-reading the data more than once, to ensure that no necessary data was left out that helps in answering the study question, and to verify the initial coding that was written and approved. Converting codes into themes or categories or category depends on its connection to the study question, and represents a typical response or a specific meaning. This stage ends with a table that identifies the candidate themes and categories. Reviewing the extracted themes in full to ensure the quality of the analysis. In this step, themes were transferred under specific categories and some codes that were not related to the study were deleted. Table 1. Shows coding book for themes and codes

Table (1): Coding book.

Theme	Codes	Example
Redesigning assessment tools	Learner styles, learner needs	"Diversifying assessment tools"
Focusing on the learning process	Shift from endproduct, student's contribution	"To reflect on their thinking"
Tasks Redesign	Ethical concerns, multimedia tools usage	"That student work is assessed holistically
Creating meaningful reviews	Critical thinking Apply knowledge	"Paying attention to real-world scenarios"

Results

Three basic themes emerged from data with

Redesigning assessment tools

Due to rapid advances in generative AI, college instructors may need to redefine and adapt their assessment strategy. For example, ChatGPT can solve hard questions quickly. 75% of participants agreed that Chatgpt can solve hard questions,

A participant suggested "we can rethink of a sound testing strategy by taking into account the goal of the assessment. Are we targeting the level of mastery learning? or our goal to focus on tracking progress and providing feedback throughout the learning process.

Another participant added" to design a reliable assessment while using generative AI, I would focus on diversifying assessment tools that fit wide range of students like special needs students by identifying learning styles.

Another participant confirmed "We should use authentic assessment that focuses on critical thinking, collaborative learning, and project-based learning. For example, incorporating multimedia elements to present ideas they get from GAI, using real-world scenarios, and creating assignments to suit learners' interests".

Focusing on the learning process

65% of participants mentioned that they should not focus on the output, while they should ask students to submit their writing in stages.

Reflecting on thinking

A participant said "we should encourage students to reflect on their thinking while they get responses from generative AI tools to create the final product. Another participant added" it is crucial to evaluate the process of learning rather than just the end product when grading!

Another participant emphasized "engaging students with science and technology specialties is in active, experiential, and project-based learning while using GAI is a necessity".

Focusing on students' contributions

65% of participant mentioned that creating collaborative learning opportunities for students that go beyond just "busy work" is crucial since it focuses on the learning process itself.

A participant said: "I tried to include collaborative and individual presentation tasks to assess learning process and I provided criteria for evaluating design, and performance.

Tasks Redesign

Many GAI tools are threating academic integrity. 72% of participants tried to reframe aspects of the tasks to mitigate the risk of such ethical concerns.

A participant mentioned "AI-resistance is a concept may sound comforting, but it is actually a misleading goal in the age of generative AI. As GAI tools are constantly evolving, becoming more sophisticated and nuanced in their outputs. The objective to be AI-resistance today may be compromised tomorrow".

Three subthemes emerged for task redesign that includes:

Providing transparent design principles

52 % of participants mentioned that "as college instructors, we should communicate and guide students about when and not use AI to solve problems, we are responsible for explaining the rationale for the task that includes providing clear guidelines regarding the appropriate use of AI tools and techniques in the learning process"

Another participant added "I think that GAI assessment tools like ChatGPT offers higher fairness with AI algorithms due to increased transparency compared with our evaluations as teachers, we do experience emotional factors sometimes, while GAI tools minimize these influences"

Multimedia and creative design elements Integration

Activities that are typically AI created like text-based should be enriched with multimedia elements like: visual, audio, and creative design elements. Many participants added: "while using GAI students should be encouraged to

Reham Salhab

use multimedia elements, concept maps, to engage students and enhance their understanding, and express their ideas

Design collaborative learning opportunities

78% participants suggested creating collaborative learning that goes beyond simple busywork while using GAI. For example, a participant said to assess students' performance there should be a focus on their contributions and learning by promoting teamwork, communication, and student collaboration.

Another participant suggested" rubrics should be created while using GAI to assess students' creativity, design, performance while using multimedia elements which ensures that student work is assessed holistically".

Creating meaningful reviews

Encouraging critical and creative thinking is essential while designing assessment in the age of general artificial intelligence.

Validating learning outcomes

63% of participants showed that "assessments should be designed in a different way that targets group work, peer, and self-assessments". Another participant said "while using GAI assessments tools, we should be selective in paying attention to real-world scenarios that promote meaningful learning experiences for learners, for example, students should be asked to connect the information they get from GAI tools to their own lives, current events, previous experiences, relevant readings, or relevant case studies.

This strategy helps validate learning outcomes by connecting the practical relevance and application of the knowledge

Discussion

The emergence of generative artificial intelligence (Gen AI), such as ChatGPT, is challenging traditional approaches to college assessment, jeopardizing the accuracy of student evaluations. AI tools can automate tests

and assignments, potentially obscuring a student's true understanding and effort. Recognizing this issue is critical to maintaining the integrity of academic assessments. Educators must adapt their assessment strategies to ensure that they still reflect students' true learning, despite the influence of AI.

Results of this study revealed four assessment strategies: assignment redesign, focusing on learning process, tasks redesign, and creating meaningful reviews. Each strategy is dedicated to providing innovative strategies for creating engaging and effective assessments tailored to undergraduate college students.

Assessment redesign is as strategy that was suggested by college instructors since integrating assessment with a clear pedagogical strategy necessitates them to rethink of the design of instruction and assessment to ensure that learning occurs. This is due to the importance of validity of assessment since there is a need to review assessments and redesign them so that students can demonstrate their learning attainment while using Generative AI-enabled. This similar with a study conducted by Dawson *et al*. (2024) who emphasized the importance of validity of assessment while using AI tools since it identifies students who have met the academic standards of a course to an agreed level of academic performance and distinguish these from those who have not met the requirement.

Task redesign is necessary to avoid academic integrity by shifting the focus from traditional evaluation methods to more meaningful learning as projects-based learning, giving presentations, or working on groups. These methods might encourage critical thinking (Ateeq et. al, 2024).

An emerging theme was focusing on the learning process instead of the final product. This is could be interpreted due to the

importance of shifting toward engaging and interaction of the students in the learning process and make them co-producers of the knowledge and information they get from GAI tools. Similarly, Awadallah Alkouk and Khlaif (2024) asserted on not focusing on the product of the GAI-assessment and instead focusing on evaluating how students develop prompts and how they work with AI throughout their learning process.

It is crucial promote meaningful learning while assessing student understanding amid the proliferation of generative AI tools by enhancing students' intrinsic motivation to learn, use transparent design principles, where teachers clearly communicate guidelines about when students should and should not use AI to solve problems. Kaldaras *et al*. (2024) showed the importance of effectiveness that supports knowledge application by incorporating features of a learning system such as creating meaningful learning opportunities and providing timely and appropriate scaffolding and feedback.

One of the themes found is to design meaningful reviews by having students connect the information they get from generative Ai tools in their written work to their own lives, current events, previous course concepts, relevant readings, case studies, or in order to verify learning outcomes. This could be explained as creating opportunities for each student to apply their knowledge they get from different tools of GAI sharpen their skills in unique settings that is specifically designed to meet their individual needs and interests. This result is in line with Khaldaras *et al*. (2024) who found that shifting from traditional assessment that requires knowledge application instead of memorization-based assessment which in turn enhances students' skills to solve real-world problems and explain real-world phenomena around them.

Conclusions

While GAI is a powerful tool, it should not be seen as the ultimate solution to all educational challenges. The revolution that GAI represents in education is redefining the assessment landscape. It no longer focuses solely on quantifying knowledge, but has a deeper purpose: to understand and support each student's individual learning journey.

Based on the results, several strategies were mentioned by college instructors include assessment redesign, focusing on the learning process over the final product, incorporating oral assessments, modifying tasks to be more specific and context-dependent, and using GAI tools critically rather than passively. Using advanced algorithms, these systems have the capacity to modify in real time both the degree of difficulty and the content of the assessments, based on the immediate responses of the students. It is crucial to teach students to be critical of the information generated by these tools and not to accept everything produced by GAI as absolute truth, as well as to foster in them a mentality of continuous learning, where GAI is just one more tool in their educational arsenal. They are more difficult to automate and offer a more comprehensive view of student skills. However, beyond technological fascination and futuristic visions, it is vital to address the present. College teachers, at all educational levels, are already integrating these tools into their academic routine, often without adequate training. These tools, while advanced and useful, can be a double-edged sword if a critical approach is not used. Blind trust in them can lead to misinformation or error.

In closing, it is essential to consider these assessment strategies and human-centered teaching philosophy, GAI will be able to act as a catalyst in optimizing the educational process and, consequently, assessment methods. College instructors need to develop GAI

assessment literacy while using GAI based assessment.

Recommendations

For responsible and effective implementation data privacy. GAI offers unprecedented opportunities to improve and optimize assessment processes. However, its implementation in the educational environment must be done with caution and responsibility. Here are some key recommendations to ensure an effective and ethical incorporation of this technology in a university education includes: training and capacity building for educators and students. It is essential that teachers are well informed about the capabilities and limitations of GAI. This involves offering courses and workshops on how to use GAI-based tools and fostering understanding of the algorithms behind these tools, so that educators can make critical and selective use of them. But we cannot forget to educate students so that they are also aware of the benefits and dangers of GAI (García-Peñalvo et al., 2024). It is crucial to promote academic integrity among students to avoid the misuse of these tools. Moreover, the issue of fairness while using GAI in assessment could be considered during the university admissions process. For example, it is necessary to employ AI algorithms for student admissions evaluations, the algorithms could be used to assess and compare candidates based on a range of objective criteria and data indicators such as academic performance and standardized test scores. This process will minimize subjective biases, thereby avoiding teacher's emotional factors. Additionally, AI algorithms utilization in assessment processes enhances transparency by offering the evaluative criteria for students. Teachers can validate learners' responses by examining the recorded data, hence providing a transparent, and equitable certification of achievements. This level of transparency helps students to understand the basis of their assessments and

facilitates justification of the results and in turn enhances learning process effectiveness.

Limitations and future research

Limitation of this study arises from the dataset being specific to a particular university and specific region. This is supported by the findings of the Moral Machine Experiment, which highlight the significant influence of national and cultural factors on AI ethics (Awad et al., 2018), thus potentially limiting the generalizability of the results. Future research could explore the factors impacting AI- based assessment strategies, and further investigations could delve into understanding the strategies in-depth and trying to find a conceptual framework. Investigating assessment strategies empirically is a good way to support results.

Disclosure Statement

Open Access

References

– Ahmed, Z., Shanto, S. S., Rime, M. H. K., Morol, M. K., Fahad, N., Hossen, M. J., & Abdullah-Al-Jubair, M. (2024). The Generative AI Landscape in Education: Mapping the Terrain of Opportunities, Challenges and Student Perception. *IEEE Access*.

– Al Harrasi, Nasser Hamed, Salah El Din, Mohamed (2024-05-06). "Utilizing AI for Assessment, Grading, and Feedback in Higher Education". IGI Global.

– Al-Ramamne, A.-L., Abeed, M., & Sabayleh, O. (2018). Assesing the Transitional Services for Persons with Mental Disabilities from Parents and Teachers Perspective. *An-Najah University Journal for Research - B (Humanities), 32*(8), 1605–1628. https://doi.org/10.35552/0247-032-008-007

– Arslan, B., Lehman, B., Tenison, C., Sparks, J. R., López, A. A., Gu, L., & Zapata-Rivera, D. (2024). Opportunities and challenges of using generative AI to personalize educational assessment. *Frontiers in Artificial Intelligence*, 7, 1460651.

– Ateeq, A., Alzoraiki, M., Milhem, M., & Ateeq, R. A. (2024, October). Artificial intelligence in education: implications for academic integrity and the shift toward holistic assessment. In *Frontiers in education* (Vol. 9, p. 1470979). Frontiers Media SA.

– Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., … & Rahwan, I. (2018). The moralmachine experiment. *Nature, 563*(7729), 59–64.

– Awadallah Alkouk, W., & Khlaif, Z. N. (2024). AI-resistant assessments in higher education: Practical insights from faculty training workshops. In *Frontiers in Education* (Vol. 9, p. 1499495). Frontiers Media SA.

– Baidoo-Anu, D., and Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): understanding the potential benefits of ChatGPT in promoting teaching and learning. *J. AI* 7, 52–62. doi: 10.61969/jai.1337500

– Boukherouaa, E. B., Shabsigh, M. G., AlAjmi, K., Deodoro, J., Farias, A., Iskender, E. S., ... & Ravikumar, R. (2021). *Powering the digital economy: Opportunities and risks of artificial intelligence in finance*. International Monetary Fund.

– Chai, F., Ma, J., Wang, Y., Zhu, J., & Han, T. (2024). Grading by AI makes me feel fairer? How different evaluators affect college students' perception of fairness. *Frontiers in Psychology*, 15, 1221177.

– Cheng, L., Croteau, E., Baral, S., Heffernan, C., and Heffernan, N. (2024). Facilitating student learning with a chatbot in an online math learning platform. *J. Educ. Comput. Res.* 62, 907–937. doi: 10.1177/07356331241226592

– Crawford J, Cowling M, Allen K. Leadership is needed for ethical ChatGPT: character, assessment, and learning using

artificial intelligence (AI). J. Univ. Teach. Learn. Pract. 2023;20(3). https://doi.org/10.53761/1.20.3.02.

– Dawson, P., Bearman, M., Dollinger, M., & Boud, D. (2024). Validity matters more than cheating. *Assessment & Evaluation in Higher Education*, 1–12. https://doi.org/10.1080/02602938.2024.2386662

– D'Agostino, T. J. (2023). Examination reform for higher order thinking: A case study of assessment-driven reform in Uganda. *International Journal of Educational Development*, *103*, 102918.

– Elmourabit, Z., Retbi, A., & El Faddouli, N. E. (2024, October). The Impact of Generative Artificial Intelligence on Education: A Comparative Study. In *Proceedings of the 23rd European Conference on e-Learning*. Academic Conferences International.

– Emmel, N. (2013). Sampling and choosing cases in qualitative research: A realist approach. Sage Publications. variation samplings

– Foung, D., Lin, L., & Chen, J. (2024). Reinventing Assessments with ChatGPT and Other Online Tools: Opportunities for GenAI-empowered Assessment Practices. *Computers and Education: Artificial Intelligence*, 100250.

– Flodén, J. (2025). Grading exams using large language models: A comparison between human and AI grading of exams in higher education using ChatGPT. *British educational research journal*, *51*(1), 201-224.

– Flores-Vivar, JM, & García-Peñalvo, FJ (2023). Reflections on the ethics, potential and challenges of Artificial Intelligence in the framework of Quality Education (SDG4). Comunicar, 31(74), 37-47. https://doi.org/10.3916/C74-2023-03

– García-Peñalvo, F.J., Corell, A., Abella-Garcia, V., & Grande-de-Prado, M. (2020). Online assessment in higher education in times of COVID-19. Education in the Knowledge Society, 21, Article 12. https://doi.org/10.14201/eks.23013

– García-Peñalvo, F. J. (2024). How general artificial intelligence affects assessment processes. Cuadernos de Pedagogía(549).

– Gökoğlu, S. (2024). "Challenges and limitations of generative AI in education," in *Transforming education with generative AI*. ed. N. Gunsel (GI Global), 158–181.

– Jin, Y., Yan, L., Echeverria, V., Gašević, D., & Martinez-Maldonado, R. (2025). Generative AI in higher education: A global perspective of institutional adoption policies and guidelines. *Computers and Education: Artificial Intelligence*, *8*, 100348.

– Jawabreh, M., & Itmazi, J. (2025). Efficiency of Artificial Intelligence in Answering Questions of Palestinian General Secondary Education Exam (Tawjihi) in Arabic Curricula: Case of (ChatGPT) and (Gemini). *An-Najah University Journal for Research - B (Humanities), 39*(8), 551–570. https://doi.org/10.35552/0247.39.8.2442

– Kabal, O., Harzallah, M., Guillet, F., & Ichise, R. (2024). Enhancing Domain-Independent Knowledge Graph Construction through OpenIE Cleaning and LLMs Validation. *Procedia Computer Science*, *246*, 2617-2626.

– Khlaif, Z. N., Ayyoub, A., Hamamra, B., Bensalem, E., Mitwally, M. A., Ayyoub, A., ... & Shadid, F. (2024). University teachers' views on the adoption and integration of generative AI tools for student assessment in higher education. *Education Sciences*, *14*(10), 1090.

– Khlaif, Z., Odeh, A., & Bsharat, T. R. (2024). Generative AI-powered adaptive assessment. In *Power of Persuasive*

Educational Technologies in Enhancing Learning (pp. 157-176). IGI Global.

– Khlaif, Z. N., Alkouk, W. A., Salama, N., & Abu Eideh, B. (2025). Redesigning Assessments for AI-Enhanced Learning: A Framework for Educators in the Generative AI Era. *Education Sciences*, *15*(2), 174.

– Kolade, O., Owoseni, A., & Egbetokun, A. (2024). Is AI changing learning and assessment as we know it? Evidence from a ChatGPT experiment and a conceptual framework. *Heliyon*, *10*(4).

– Koseda, E., Cohen, I. K., McIntosh, B., & Cooper, J. (2025). Internationalisation and digital transformation in HEIs: The impact of education 4.0 on teaching, learning and assessment. *Policy Futures in Education*, *23*(1), 1-9.

– Kumar, T., Kait, R., Ankita, Malik, A. (2024). The Role of Generative Artificial Intelligence (GAI) in Education: A Detailed Review for Enhanced Learning Experiences. In: Shukla, B., Murthy, B.K., Hasteer, N., Kaur, H., Van Belle, JP. (eds) Intelligent IT Solutions for Sustainability in Industry 5.0 Paradigm. ICEIL 2023. Lecture Notes in Electrical Engineering, vol 1185. Springer, Singapore. https://doi.org/10.1007/978-981-97-1682-1_17

– Levitt, H. M., Bamberg, M., Creswell, J. W., Frost, D. M., Josselson, R., & Suárez-Orozco, C. (2018). Journal article reporting standards for qualitative primary, qualitative meta-analytic, and mixed methods research in psychology: The APA Publications and Communications Board task force report. *American Psychologist*, 73(1), 26–46. https://doi.org /10.1037/amp0000151

– Liu, Y., Han, T., Ma, S., Zhang, J., Yang, Y., Tian, J., He, H., Li, A., He, M., Liu, Z., Wu, Z., Zhao, L., Zhu, D., Li, X., Qiang, N., Shen, D., Liu, T., & Ge, B. (2023). Summary of ChatGPT-Related research and perspective towards the future of large language models. *Meta-Radiology,* 1(2), Article 100017. https://doi.org/10.1016/j.metrad.2023.100017

– Madland, C., Irvine, V., DeLuca, C., & Bulut, O. (2024). Developing the Technology-Integrated Assessment Framework. *The Open/Technology in Education, Society, and Scholarship Association Journal*, *4*(1), 1-19.

– Mao, J., Chen, B., and Liu, J. C. (2024). Generative artificial intelligence in education and its implications for assessment. TechTrends 68, 58–66. doi: 10.1007/s11528-023-00911-4

– Marengo, A. (2024). Navigating the nexus of AI and IoT: A comprehensive review of data analytics and privacy paradigms. *Internet of Things*, 101318.

– McDonald, N., Johri, A., Ali, A., & Collier, A. H. (2025). Generative artificial intelligence in higher education: Evidence from an analysis of institutional policies and guidelines. *Computers in Human Behavior: Artificial Humans*, 100121.

– Memarian, B., & Doleck, T. (2024). A review of assessment for learning with artificial intelligence. *Computers in Human Behavior: Artificial Humans*, *2*(1), 100040.

– Miller, J. C., Miranda, J. P. P., & Tolentino, J. C. G. (2025). Artificial Intelligence in Physical Education: A Review. *Global Innovations in Physical Education and Health*, 37-60.

– Mohammed, C., & Rocke, S. (2024, October). WIP Post-Assessment Processes Given the Rise of Generative AI: Findings from the Literature. In *2024 IEEE Frontiers in Education Conference (FIE)* (pp. 1-4). IEEE.

- Mohajan, H. K. (2018). Qualitative Research Methodology in Social Sciences and Related Subjects. Journal of Economic Development, Environment, and People, 7, 23-48. https://doi.org/10.26458/jedep.v7i1.571

- Nedungadi, P., Tang, K. Y., & Raman, R. (2024). The Transformative Power of Generative Artificial Intelligence for achieving the sustainable development goal of Quality Education. *Sustainability*, *16*(22), 9779.

- Ng, S. H. S., Chan, H. Y., Wong, J. H. K., Sam, L., & Privitera, A. J. (2024). A Scoping Review of the Use of Generative AI in Assessment in Higher Education.

- Owoseni, A., Kolade, O., & Egbetokun, A. (2024). Applications of Generative AI in Summative Assessment. In *Generative AI in Higher Education: Innovation Strategies for Teaching and Learning* (pp. 97-122). Cham: Springer Nature Switzerland.

- Pan, F. (Ed.). (2024). *AI in Language Teaching, Learning, and Assessment*. IGI Global.

- Rezai, A. (2022). Fairness in classroom assessment: development and validation of a questionnaire. *Language Testing in Asia*, *12*(1), 17.

- Rudolph, J., Tan, S., and Tan, S. (2023). ChatGPT: bullshit spewer or the end of traditional assessments in higher education? *J. Appl. Learn. Teach.* 6. doi: 10.37074/jalt.2023.6.1.9

- Sabet, P. G. P., Zhan, S. J., & Shishehgarkhaneh, M. B. (2024). AI and Assessment Design in Engineering Disciplines. *Education Research & Perspectives*, *51*.

- Saihi, A. (2024). "A Structural equation modeling analysis of generative AI chatbots adoption among students and educators in higher education". 7(N/A).

- https://www.sciencedirect.com/science/article/pii/S2666920X24000778.

- Samala, A. D., Zhai, X., Aoki, K., Bojic, L., and Zikic, S. (2024). An in-depth review of ChatGPT's pros and cons for learning and teaching in education. Int. J. Interact. Mob. Technol. 18, 96–117. doi: 10.3991/ijim.v18i02.46509

- Saputra, I., Kurniawan, A., Yanita, M., Putri, E. Y., & Mahniza, M. (2024). The Evolution of Educational Assessment: How Artificial Intelligence is Shaping the Trends and Future of Learning Evaluation. *The Indonesian Journal of Computer Science*, *13*(6).

- Sesay, I. M. S. G. A. (2024). Redefining Student Assessments in the Age of AI-Assisted Learning Strategies for Educators. *GSJ*, *12*(8).

- Seidman, G. (2013). Self-presentation and belonging on Facebook: How personality influences social media use and motivations. *Personality and Individual Differences*, 54(3), 402–407. https://doi.org/10.1016/j.paid.2012.10.009

- Thomas, G. (2011). *How to do your case study: A guide for students and researchers*. Thousand Oaks, US: Sage Publications.

- Vetrivel, S. C., Vidhyapriya, P., & Arun, V. P. (2025). The Role of AI in Transforming Assessment Practices in Education. In *AI Applications and Strategies in Teacher Education* (pp. 43-70). IGI Global.

- Wu, H., Li, D., & Mo, X. (2025). Understanding GAI risk awareness among higher vocational education students: An AI literacy perspective. *Education and Information Technologies*, 1-32.

- .Xia, Q., Weng, X., Ouyang, F., Lin, T. J., & Chiu, T. K. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education.

International Journal of Educational Technology in Higher Education, *21*(1), 40.

– Yin, RK (2003) Case Study Research: Design and Methods. 3rd Edition, Sage, Thousand Oaks.

– Zuckerman, M., Flood, R., Tan, R. J., Kelp, N., Ecker, D. J., Menke, J., & Lockspeiser, T. (2023). ChatGPT for assessment writing. *Medical Teacher*, *45*(11), 1224-1227.