



Ensuring academic integrity in the age of ChatGPT: Rethinking exam design, assessment strategies, and ethical AI policies in higher education

Edmund De Leon Evangelista ^{1*}

 0000-0002-8050-1360

¹ Zayed University, Dubai, UNITED ARAB EMIRATES

* Corresponding author: edmund.evangelista@zu.ac.ae

Citation: Evangelista, E. D. L. (2025). Ensuring academic integrity in the age of ChatGPT: Rethinking exam design, assessment strategies, and ethical AI policies in higher education. *Contemporary Educational Technology*, 17(1), ep559. <https://doi.org/10.30935/cedtech/15775>

ARTICLE INFO

Received: 18 Sep 2024

Accepted: 6 Dec 2024

ABSTRACT

The rapid advancement of artificial intelligence (AI) technologies, particularly OpenAI's ChatGPT, has significantly impacted higher education institutions (HEIs), offering opportunities and challenges. While these tools enhance personalized learning and content generation, they threaten academic integrity, especially in assessment environments. This study systematically examines the impact of ChatGPT on academic integrity in HEIs, focusing on exam design, assessment strategies, AI detection tools, and policy frameworks. This research draws from current literature and expert recommendations to identify practical approaches for developing assessments that foster critical thinking and deep cognitive engagement, making them less susceptible to AI-generated content. Key areas explored include the creation of complex, analytical exam formats, deploying advanced AI detection software to counter AI-assisted cheating, and formulating institutional policies that promote the ethical use of AI. This comprehensive framework aims to equip educators and administrators with practical strategies to preserve academic standards while harnessing the potential benefits of AI, ensuring the continued validity of assessments in the AI-driven educational landscape.

Keywords: ChatGPT, academic integrity, higher education, exam design, AI detection, assessment strategies, AI policies

INTRODUCTION

The rapid proliferation of artificial intelligence (AI) technologies, particularly OpenAI's ChatGPT, has significantly transformed the educational landscape, reshaping how students learn and how educators design assessments. ChatGPT, a large language model (LLM) capable of generating human-like text based on user inputs, offers numerous advantages in higher education. It enhances personalized learning, automates routine tasks, and provides students instant feedback, tutoring, and support across various academic disciplines (Halaweh, 2023). However, the increasing use of AI also raises substantial concerns about academic integrity, as these tools enable students to generate assignments, essays, and problem solutions with minimal effort or understanding, thus circumventing traditional learning and assessment processes (Imran & Almusharraf, 2023; Lo, 2023).

Integrating AI into education requires educators and institutions to rethink how they design assessments and uphold academic integrity in higher education. Bloom's Taxonomy offers a valuable framework for this rethinking, as it categorizes cognitive learning objectives from lower-order thinking skills (such as remembering and understanding) to higher-order skills (such as analyzing, evaluating, and creating). Traditional assessments, which often emphasize lower-order cognitive skills, are particularly vulnerable to AI-generated content, as tools like ChatGPT can easily replicate factual recall and basic comprehension tasks. In contrast, assessments that focus on higher-order cognitive skills—requiring students to analyze complex

problems, evaluate different perspectives, and create original work—pose a more significant challenge for AI to mimic. Therefore, reorienting assessments toward these higher levels of Bloom's Taxonomy fosters deeper learning and strengthens academic integrity by making it harder for AI-generated responses to pass undetected.

Beyond cognitive skills, educators and institutions must also consider the ethical implications of AI usage in education. The ethical framework for responsible AI use, rooted in principles of fairness, transparency, and accountability, is crucial in ensuring that AI enhances learning without compromising the integrity of the educational process. As educators and institutions adopt AI, they must establish clear guidelines that define the acceptable and ethical use of AI tools like ChatGPT. Such guidelines should emphasize the development of AI literacy among students and staff, ensuring that all parties understand how to use AI responsibly and the potential consequences of misuse. It includes a focus on fairness, as unchecked AI use can create inequities, particularly for students who may not have access to AI tools or who are not adequately trained in their ethical use (Floridi, 2023). Additionally, transparency in AI policies is vital—institutions must be clear about how AI-generated content is detected, evaluated, and penalized to maintain trust and uphold academic standards (Cotton et al., 2023).

The misuse of AI tools for academic dishonesty also raises important questions about the responsibilities of educators in mitigating these risks. Higher education institutions (HEIs) must design robust assessment frameworks beyond traditional approaches, incorporating ethical AI practices and fostering a culture of academic integrity. This study systematically reviews best practices for exam design and assessment strategies in the age of AI, focusing on how HEIs can create assessments that demand higher-order cognitive engagement and reduce reliance on AI-generated content. Furthermore, the study investigates the role of AI detection tools and institutional policies in maintaining academic standards.

The key areas explored include the following:

1. **Exam design:** Developing assessments that emphasize complex reasoning, real-world problem-solving, and the application of knowledge in unique contexts, all of which require students to engage in higher-order cognitive processes.
2. **Assessment strategies:** Implementing innovative evaluation techniques that measure students' deeper understanding and critical thinking skills while minimizing opportunities for AI-assisted cheating.
3. **AI detection tools:** Using advanced software to identify AI-generated content, ensuring the authenticity of student work.
4. **Policy and training:** Establishing ethical policies and conducting training programs for both students and educators to promote the responsible use of AI while safeguarding academic integrity.

In this AI-driven era, ensuring academic integrity goes beyond simply policing AI use—it requires a comprehensive approach that integrates higher-order cognitive skills, ethical considerations, and effective assessment strategies. This study provides a detailed framework for educators and institutions to adapt their practices in response to the opportunities and challenges posed by AI, aiming to uphold the value and credibility of academic credentials in a rapidly evolving technological landscape.

ACADEMIC INTEGRITY AND RELATED WORKS

Incorporating AI technologies like ChatGPT into educational environments has sparked a heated debate about academic integrity. ChatGPT, designed by OpenAI, utilizes natural language processing to render human-like text, presenting concerns about its potential misuse in academic settings. This review summarizes the literature on ChatGPT's implications for academic integrity and investigates best practices for enforcing robust exams and assessment strategies in higher education.

ChatGPT and Academic Integrity

AI's role in education has grown significantly, with benefits including personalized learning experiences, integrated educational content, and adaptability to student needs. Montenegro-Rueda et al. (2023) contend that employing ChatGPT in HEIs has sparked curiosity because it can enhance students' learning experiences,

deliver instantaneous feedback, and manage personalized needs. On the contrary, its usage has introduced recent concerns about academic integrity.

Academic integrity is vital to HEIs, as it guarantees the credibility of qualifications and steadfastness of scholarly work. It ensures that academic staff, students, and researchers follow moral and professional norms, values, and principles in teaching, research, and scholarship (Lancaster, 2024). Hence, the intellectual achievements of others are respected while aiming for academic excellence. Even so, ChatGPT can be used to perpetrate academic fraud, and their misuse undermines the educational process and devalues genuine student effort. In addition, such tools may provide a platform for new types of serious academic misconduct that are difficult to detect and prove (Cotton et al., 2023).

Researchers have observed ChatGPT's ability to generate refined responses indiscernible from student-generated content (Floridi, 2023). This ability enables students to avoid traditional plagiarism detection methods. Educators are growingly concerned that students will employ ChatGPT to create written assignments (Halaweh, 2023; Meyer et al., 2023). It has been proven to render reports in seconds uncaught by plagiarism detectors. In the same way, Talan and Kalinkara (2023) used a 40-item multiple-choice exam in their study to compare ChatGPT's performance to undergraduate students in a Turkish state university's anatomy course. The findings showed that ChatGPT performed better than the students. Hence, it is suggested that instantaneous action be taken to revise assessment procedures and institutional guidelines in schools and universities (Lo, 2023).

Related Works

AI's role in education has grown dramatically, providing transformative benefits such as personalizing learning experiences for individual students, seamlessly integrating educational materials, and dynamically adapting to diverse learning needs. These innovations create more engaging and responsive learning environments, permitting educators to address each student's strengths and challenges. Even so, Klyshbekova and Abbott (2024) revealed that ChatGPT could behave like a magic wand to accomplish tasks quickly and competently, making it nearly impossible to tell the difference between students' words and those of the AI tool. Similarly, Vargas-Mutillo et al. (2023) observed that students who primarily depend on ChatGPT to finish their homework risk developing a dependency that impedes their intellectual development. This dependency hinders their capacity to become interactive, objective, and imaginative.

The prospect of ChatGPT compromising academic integrity is a typical highlight in recent literature. Bukar et al. (2024) and Rahimi et al. (2024) conducted studies demonstrating diverse forms of academic dishonesty enabled by technology, ranging from plagiarism to contract cheating. Contract cheating is the act of exchanging money for the provision of custom-written coursework. It is also known as assignment commissioning or outsourcing. Equally, McIlwraith et al. (2023) examined GPT-3's capability to construct credible academic texts and to prove that students use it to cheat. Experimental results revealed that GPT-3 can be an effective collaborator in paper design. Moreover, the study of Mapletoft et al. (2024) demonstrated that GPT-4 could generate high-quality and credible academic content, making it challenging to identify work that is generated by AI. They emphasized how GPT-4 can produce thorough essays or responses and replicate human writing styles accurately.

As a result, numerous studies concentrated on offering solutions to reduce ChatGPT's misuse while upholding academic integrity. Cotton et al. (2023) proposed that faculty can assess a student's comprehension of the subject matter using automated and manual assessment methods. This suggestion is in addition to providing precise and thorough instructions for assignments and marking them using a rubric. In the same way, Plata et al. (2023) emphasized that students must understand the value of academic integrity and ethical behavior, how to avoid academic dishonesty, and the repercussions of academic misconduct when using AI. Additionally, Sullivan et al. (2023) concentrated on strategies for reorganizing assignments or tasks that ChatGPT cannot complete. They also identified scenarios where students might be penalized or fail because they used ChatGPT.

Two case studies from Stanford University and Duke University provide valuable insights into how HEIs redesign assessments and enforce policies to maintain academic standards, providing a more practical perspective on how they respond to the academic integrity challenges posed by AI tools such as ChatGPT. For

example, Stanford University changed its assessment design from traditional essays to more interactive, project-based evaluations. This method demands students to apply theoretical knowledge to real-world problems, making it difficult for AI tools such as ChatGPT to generate appropriate responses. Similarly, Duke University established an "AI Ethics Committee" to regularly update institutional policies on using AI in coursework, ensuring that students and faculty know ethical boundaries (Plata et al., 2023).

Equally, the University of Sydney implemented oral examinations and group projects in response to the rise of AI tools. These formats foster critical thinking and collaboration, lowering the risk of AI misuse while promoting deeper learning (Cotton et al., 2023). These case studies show that when HEIs take proactive steps to redesign assessments and establish comprehensive policies, they can effectively safeguard academic integrity while incorporating new technologies into the learning environment.

ChatGPT's potential to render a wide range of textual content quickly raises the alarm, forcing educators to devise new approaches for detecting and preventing academic misconduct. Even before LLMs, there was a prolonged struggle over academic integrity issues. Yet, this struggle was made worse by the absence of anti-cheating and anti-plagiarism policies (Plata et al., 2023). While the aforementioned studies highlighted specific issues related to ChatGPT's usage, only a handful of HEIs developed a comprehensive framework to adapt to this tool's challenges. A more comprehensive systematic review is required to guarantee the preservation of academic integrity, maximize ChatGPT's advantages, and minimize its disadvantages. This study aimed to disseminate knowledge on how to rethink exams, use technology wisely, encourage ethical use, and promote policies to reduce the negative impact of AI tools. University administrators can use the insights gained from this study to formulate a framework for institutional policy against ChatGPT's misuse and eventually construct robust exams and assessments that sustain academic integrity.

METHODOLOGY

This systematic review aimed to synthesize the best approaches for constructing robust exams in higher education to sustain academic integrity in the era of AI language models, specifically ChatGPT. The review pursued the preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines (Page et al., 2021) to guarantee a comprehensive and transparent reporting process. This study can systematically identify, screen, and assess relevant studies, including high-quality and reliable sources, by following PRISMA. This process improves the research's reproducibility and ensures that the methodology is rigorous, thorough, and bias-free. PRISMA is particularly useful in defining explicit inclusion and exclusion criteria, documenting the selection process, and demonstrating the findings in a flow diagram for clarity.

Search Strategies and Limits

The study began with a review of the literature and case studies on ChatGPT's application in higher education, specifically about exam design, proctoring techniques, academic policies, and tools for upholding academic integrity. Prominent databases, specifically Scopus, Google Scholar, and JSTOR, were methodically reviewed for this process utilizing different combinations of the subsequent search terms in their title, abstract, or keywords:

("chatgpt" OR "gpt") AND ("higher education" OR "hei*") AND ("exam" OR "assessment") AND ("policy" OR "guidelines" OR "rules" OR "design" OR "practices" OR "proctor*" OR "strategy") AND ("academic integrity" OR "integrity") AND ("cheat*" OR "plagiarism").*

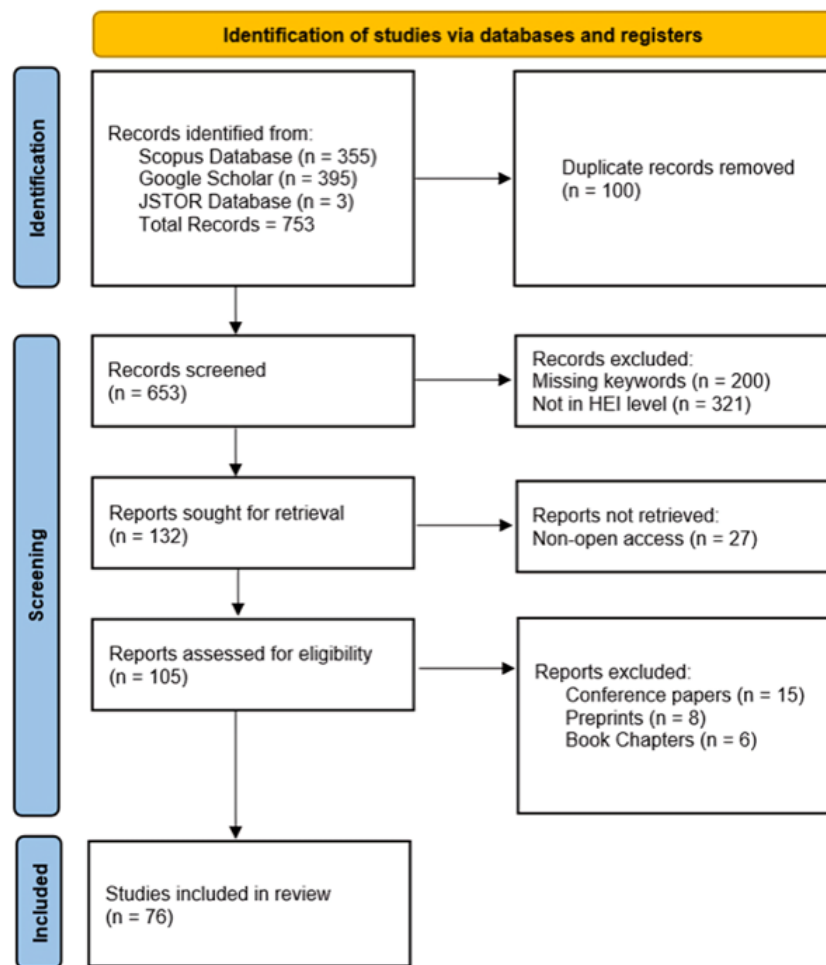
The search for articles took place in July and August of 2024. The results contained 753 records initially identified.

Inclusion and Exclusion Criteria

The search for academic articles was conducted using inclusion and exclusion criteria to guarantee the quality and applicability of the chosen studies. Publications from 2022 to the present, concentrating on higher education, discussing ChatGPT's effect on academic integrity, and offering methods or best practices for preserving an exam's integrity were all included in the inclusion criteria. This study selected this period to include the most current and pertinent literature during data retrieval, starting from the year ChatGPT was launched. Only peer-reviewed English-language articles have been considered.

Table 1. Inclusion and exclusion criteria

	Inclusion criteria	Exclusion criteria
Article topic	Papers presenting strategies or best practices for upholding academic integrity in ChatGPT	Papers not presenting strategies or best practices for upholding academic integrity in ChatGPT
Academic level	Higher education	Non-higher education
Article type	Published in peer-reviewed journal, open access	Conference proceedings, not published in peer-reviewed journal, non-open access, preprints, editorials, book chapters
Period published	Published between 2022 - present	Published before 2022
Language	English	Non-English

**Figure 1.** PRISMA diagram of article selection (Source: <https://prisma-statement.org/>)

On the other hand, studies unrelated to higher education, those limited to K-12 education, publications devoid of theoretical analysis or empirical support, and articles not available in English were excluded by the exclusion criteria. Articles in any conference proceedings, including non-peer-reviewed journals, were also excluded. The goal of this meticulous selection procedure was to produce a thorough and targeted synthesis of the current state of the field's knowledge. Any literature reviews found were used as background information but excluded from the synthesis to eliminate redundancy. **Table 1** details the specific inclusion and exclusion criteria for selecting articles.

Search Outcomes

The study performed a selection process to identify and remove duplicate publications from the selected databases, excluding 100 duplicates from the 753 initially collected documents as outlined in the PRISMA flow diagram (**Figure 1**).

Table 2. Challenges posed by AI in exam design

Challenge	Description	Implication for exam design	Authors
Factual recall vulnerability	AI tools like ChatGPT easily handle fact-based questions.	Exams should move away from factual recall questions towards those assessing deeper understanding.	Gamage et al. (2023); Susnjak and McIntosh (2024)
Lack of student engagement	Traditional exams may not sufficiently engage students in active learning.	Integrate real-world, interactive components into exams to enhance engagement.	Ali et al. (2024); Güner et al. (2024)
Ease of AI cheating	Simple, predictable questions allow AI to generate correct answers.	Develop complex, layered, and non-repetitive questions to make AI-generated responses less effective.	Moorhouse et al. (2023); Richards et al. (2024)

Table 3. Alternative assessment formats

Assessment format	Description	Benefits	Authors
Oral exams	Assessments based on spoken responses rather than written ones.	Harder for AI to assist; tests real-time reasoning and knowledge application.	Gamage et al. (2023); Susnjak and McIntosh (2024)
Project-based learning	Tasks that involve long-term engagement with real-world problems or scenarios.	Encourages critical thinking and creativity; more difficult for AI to replicate.	Chaudhry et al. (2023); Moorhouse et al. (2023)
Performance-based assessments	Students demonstrate their learning through tasks like presentations or real-world problem-solving exercises.	Requires personal insight and critical reasoning, areas where AI performs less effectively.	Richards et al. (2024); Rudolph et al. (2024)

The screening process helped remove 200 articles that did not contain the required keywords and 321 articles that did not cover the study's HEI level requirement. Additionally, from the 132 remaining articles sought for retrieval, another 27 were excluded because it is non-open access. In the last screening stage, out of 105 articles assessed for eligibility, 29 were removed further for being classified as conference papers, preprints, or book chapters. Finally, 76 articles were selected for inclusion in the literature search. The PRISMA guidelines guaranteed that only trustworthy and high-quality data were selected for this investigation.

RESULTS

After assessing the 76 articles obtained through the PRISMA flowchart, the literature review identified the following primary themes related to ensuring academic integrity with ChatGPT.

Effective Exam Design

Various studies emphasize the need to reassess traditional exam designs due to the increasing prevalence of AI tools like ChatGPT. For instance, traditional assessments such as essays and multiple-choice questions (MCQs) are seen as vulnerable to AI misuse. To address this, some studies recommend shifting toward alternative formats like oral exams, open-ended questions, and performance-based evaluations, which require more profound understanding and personal insights from students, reducing the risk of AI exploitation (Gamage et al., 2023; Susnjak & McIntosh, 2024). Articles also suggest incorporating real-world problem-solving tasks and scenarios to challenge students beyond factual recall, which AI can easily handle (Chaudhry et al., 2023; Rudolph et al., 2024).

Table 2 shows that traditional evaluation formats, particularly those that depend on factual recall, are becoming more vulnerable to AI misuse. Since ChatGPT and other AI tools can effortlessly manage MCQs and essay-based assessments, educators have difficulty upholding academic integrity. The findings highlight the need to shift toward more complicated, interactive, and engaging evaluations, like oral exams, problem-solving exercises, and multimodal questions, to ensure authentic student performance and reduce the impact of AI-assisted cheating.

As presented in **Table 3**, the findings suggest that various forms of evaluation, like performance-based assessments, project-based learning, and oral exams, provide more effective ways to decrease AI misuse. These formats promote critical thinking, creativity, and quick problem-solving skills that are difficult for AI systems to imitate. These assessment types reduce the possibility of AI-generated responses and help

Table 4. Strategies for reducing AI misuse in exams

Strategy	Details	Implementation	Authors
Real-world problem-solving	Integrating tasks that require students to apply theoretical knowledge to practical, real-world challenges.	Use case studies or simulations that reflect current industry practices.	Chaudhry et al. (2023); Moorhouse et al. (2023)
Interactive feedback mechanisms	Continuous feedback during assessments helps students engage more deeply in the learning process.	Design assignments that include checkpoints with instructor feedback throughout.	Gamage et al. (2023); Susnjak and McIntosh (2024)
Multimodal question integration	Incorporating visuals, videos, or other media to create multi-dimensional exam questions that AI struggles to solve.	Use a combination of diagrams and text-based questions to assess broader competencies.	Richards et al. (2024); Rudolph et al. (2024)

Table 5. Exam redesign principles in the era of AI

Principle	Key focus	Objective	Authors
Personalization	Create assessments that reflect individual student experiences.	Prevent students from generating AI-assisted answers by making the assessment unique to each learner.	Chaudhry et al. (2023); Susnjak and McIntosh (2024)
Process over product	Assess the learning process rather than just the final product.	Ensure students cannot use AI to shortcut the process by evaluating ongoing engagement and understanding.	Gamage et al. (2023); Moorhouse et al. (2023)
Emphasizing creativity	Focus on tasks that require novel thinking, creative problem-solving, and innovation.	Make AI less useful by assessing non-formulaic and creative responses.	Richards et al. (2024); Rudolph et al. (2024)

guarantee authentic student performance by underlining student engagement and demanding deeper but more personal insights.

The findings emphasize that oral exams stimulate critical thinking and deeper cognitive engagement by requiring students to articulate and defend their reasoning in real-time. This methodology ensures that students demonstrate authentic understanding, a crucial metric that AI-generated responses cannot mimic. For example, in a humanities class, oral exams might involve critiquing a literary text, while in engineering, students could explain the mechanics behind a design project. These scenarios highlight the versatility of oral exams across disciplines and their potential to uphold academic integrity.

The findings, as shown in [Table 4](#), demonstrate that constructing exams highlighting real-world problem-solving, interactive feedback mechanisms, and multimodal questions are effective techniques for reducing the misuse of AI in exams. These strategies require students to demonstrate original thinking, continuous engagement, and the capability to combine various information types, making it difficult for AI to generate appropriate or correct answers. By shifting emphasis from simple recall to higher-order thinking tasks, these approaches help safeguard academic integrity from AI misuse.

Moreover, various studies emphasized the need to forge new exam design principles in the era of AI to generate actionable insights for educators and institutions looking to enhance assessment design in the context of AI's growing role in education. The data in [Table 5](#) illustrates that effective exam redesign in the AI era involves concentrating on personalization, process-oriented assessments, and fostering creativity. By mandating students to commit deeply to the material, exhibit original thought, and demonstrate their learning process, these principles seek to reduce the misuse of AI. Customizing tests and focusing on creativity make it challenging for AI to generate accurate results. In contrast, process-oriented assessments guarantee students' active participation throughout the test rather than concentrating only on the outcome.

The findings in this theme highlight the importance of rethinking traditional exam formats in light of the growing capabilities of AI tools such as ChatGPT. Approaches such as project-based learning, oral exams, and open-ended questions have emerged as more effective methods for assessing critical thinking and creativity, which AI cannot easily replicate. These findings highlight the importance of developing assessments that push students beyond habitual memorization and encourage deeper engagement with the material. Educators must persist in looking for innovative assessment methods that maintain academic integrity while adapting to the changing technological landscape.

Table 6. Summary of key findings on innovative assessment strategies

Strategy	Description	Impact/benefits	Authors
Personalized learning	Using AI to provide tailored learning experiences based on individual student needs.	Offers adaptive feedback and allows for more individualized assessments that cater to diverse learning styles.	Dai et al. (2023); Güner et al. (2024)
Collaborative projects	Encouraging group discussions and teamwork-based assessments.	Promotes critical thinking, creativity, and real-world problem-solving, reducing reliance on AI-generated responses.	Chaudhry et al. (2023); Moorhouse et al. (2023)
Real-time feedback and adaptive quizzes	Incorporating AI to deliver instant feedback and dynamic assessments that adjust to student performance.	Enhances engagement and allows for continuous learning, helping students improve progressively based on immediate feedback.	Gamage et al. (2023); Rudolph et al. (2024)
Multimodal assessments	Using diverse media (e.g., images, videos, diagrams) in assessments to challenge students in multiple ways.	Difficult for AI to handle, encourages deeper cognitive engagement by testing multiple competencies beyond text-based answers.	Richards et al. (2024); Susnjak and McIntosh (2024)
Critical thinking and open-ended tasks	Designing assessments that focus on creativity, analysis, and problem-solving rather than factual recall.	Encourages higher-order thinking skills, areas where AI tools struggle, ensuring more authentic student engagement.	Chaudhry et al. (2023); Gamage et al. (2023)

Innovative Assessment Strategies

In addition to redesigning exams, innovative assessment strategies that challenge students' critical thinking and creativity, two areas where AI-generated content is less effective - must be investigated. Implementing these strategies is essential to adapt educational assessments in response to the growing influence of AI tools like ChatGPT.

The results in **Table 6** reveal that innovative assessment strategies, with a strong emphasis on **personalization**, **collaboration**, and most importantly, **real-time feedback**, can significantly enhance student engagement and reduce AI misuse. Personalized and adaptive learning tools provide tailored assessments, while collaborative projects encourage critical thinking and creativity.

Project-based learning encourages critical thinking by engaging students in complex, real-world scenarios that require applying theoretical knowledge creatively. For instance, engineering students might work collaboratively to design sustainable urban solutions, while business students could develop strategic market plans for a hypothetical product. These tasks minimize the risk of AI misuse and encourage skills such as collaboration, adaptability, and innovation, which are critical for workplace success.

Moreover, real-time feedback, in particular, can provide students with immediate insights into their performance, fostering a sense of urgency and continuous improvement. **Multimodal assessments**, which combine various media and open-ended tasks that emphasize critical thinking, promote deeper cognitive engagement, making it difficult for AI to generate accurate responses. These strategies collectively sustain more authentic learning experiences.

Similarly, the additional approaches emphasized in different research studies below confirm that assessments stay impactful and pertinent in an educational environment influenced by AI. These approaches could boost student involvement, foster genuine learning, and reinforce academic honesty by emphasizing analytical thinking, practical, real-life use, and continuous feedback, thereby increasing the resilience of assessments against AI malpractice and better equipping students for upcoming obstacles.

1. **Integration of AI into assessments:** Encouraging students to critique or improve AI-generated content helps them engage more critically with both the material and AI itself. This deepens their understanding of the subject and builds awareness of AI's limitations, turning AI into a learning tool rather than a source of shortcuts (Dai et al., 2023; Gamage et al., 2023).
2. **Formative assessments:** Low-stakes, continuous assessments provide ongoing feedback, allowing students to reflect on their learning progress. This reduces reliance on one-time, high-stakes exams that AI can easily exploit, promoting consistent engagement and personal growth (Rudolph et al., 2024; Susnjak & McIntosh, 2024).

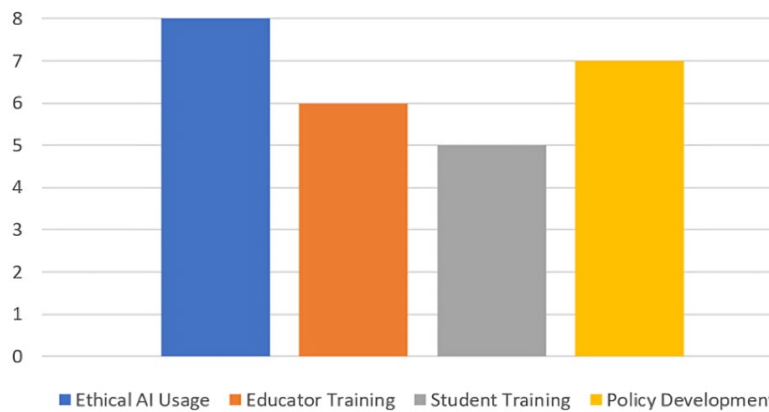


Figure 2. Frequency of innovative assessment strategies mentioned, based on the findings of the 76 articles systematically reviewed in this study

Table 7. List of advanced AI detection tools

Tool/approach	Description	Impact/benefits	Authors
AI text detection software	Software designed to detect AI-generated content in student submissions (e.g., Turnitin AI detection & GPTZero).	Identifies instances of AI-assisted cheating by analyzing writing patterns and content generation styles.	Gonsalves (2023); Moorhouse et al. (2023)
Plagiarism detection enhanced by AI	Using traditional plagiarism detection software that has been adapted to recognize AI-generated text.	Helps detect both traditional plagiarism and AI-generated content, ensuring broader coverage of academic misconduct.	Gamage et al. (2023); Rudolph et al. (2024)
Linguistic fingerprinting	Analyzing writing style and linguistic features to detect inconsistencies in student submissions potentially caused by AI.	Compares writing patterns to detect deviations that suggest AI involvement, especially in more personalized assignments.	Gamage et al. (2023); Susnjak and McIntosh (2024)
Advanced algorithms and machine learning models	Utilizing sophisticated algorithms to recognize AI-generated patterns in student work.	Offers more accurate detection as AI tools evolve, providing better protection against increasingly sophisticated AI models.	Chaudhry et al. (2023); Richards et al. (2024)

- 3. Real-world applications:** Using practical, real-world scenarios in assessments requires students to apply theoretical knowledge in ways that AI cannot easily replicate. This fosters problem-solving, creativity, and critical thinking, ensuring that students are better prepared for real-world challenges while minimizing the potential for AI misuse (Chaudhry et al., 2023; Moorhouse et al., 2023).

Additionally, the graph in **Figure 2** demonstrates that project-based learning is the most commonly mentioned innovative assessment strategy, followed by peer assessments and oral exams. Self-reflective assessments and open-book exams are less frequently mentioned, but they are still regarded as important in addressing the challenges posed by AI in education. The findings in this theme demonstrate the effectiveness of project-based learning, oral exams, and peer assessments in promoting critical thinking and creativity, areas where AI tools like ChatGPT fall short. These strategies increase student engagement and play an essential role in preserving academic integrity in the face of rapidly evolving AI technologies.

Advanced AI Detection Tools

While redesigning exams is an important step, it is also critical to develop and improve AI detection tools that can assist educators in identifying AI-generated content. The tools mentioned in the results section show promise, but they must be constantly updated to keep up with advances in AI technology. Leveraging advanced AI detection tools is crucial for maintaining academic integrity by ensuring that assessments accurately reflect students' genuine efforts.

As outlined in **Table 7**, the findings reveal that diverse AI detection tools, including text recognition software, AI-enhanced plagiarism detection, and linguistic profiling, are being developed and utilized to identify AI-generated content in student work.

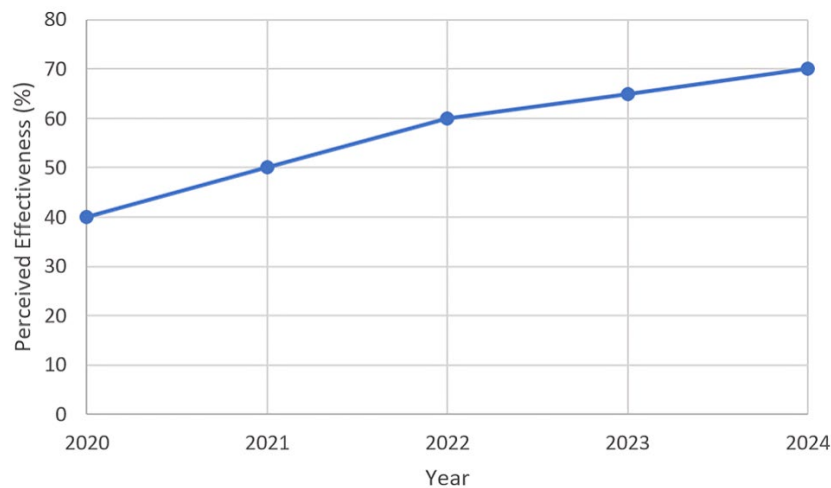


Figure 3. Effectiveness of AI detection tools over time, based on the findings of the 76 articles systematically reviewed in this study

Table 8. Advanced AI detection tools and their effectiveness in educational settings

AI detection tool	Description	Effectiveness	Limitations
GPTZero	AI detection tool designed to identify AI-generated text (Nikolic et al., 2023).	Moderate effectiveness in identifying AI-generated content in specific contexts.	Early-stage development, still requires refinement for more accurate detection.
Turnitin’s AI detection feature	Integrated AI detection tool within Turnitin, used to detect plagiarism and AI-generated content (Gonsalves, 2023).	Claims high accuracy in detecting AI-generated text, particularly useful for educational settings.	Concerns about privacy and the accuracy of detection, especially with heavily paraphrased or human-edited AI content.
DetectGPT	A tool developed to specifically detect text generated by GPT models (Richards et al., 2024).	Effective at identifying patterns typical of AI-generated content.	Struggles with detecting well-edited or paraphrased AI content, prone to false positives.
ZeroGPT	AI detection tool focusing on identifying AI-written content (Klyshbekova et al., 2024).	Useful for detecting early-stage AI-generated responses.	Can be circumvented by more sophisticated AI models or edited text, requiring ongoing updates.
GPT-2 detector	Detection model specifically aimed at identifying content generated by earlier versions of GPT models (Richards et al., 2024).	Limited success with newer models like GPT-4 due to advancements in AI sophistication.	Not reliable for detecting content from more advanced AI models, leading to inconsistent results.

Furthermore, sophisticated algorithms and machine learning models are being deployed to enhance accuracy as AI tools progress, offering stronger protections against academic dishonesty. These tools bolster institutions’ capability to maintain academic integrity in an AI-centric educational setting.

The graph depicting the effectiveness of AI detection tools over time, as illustrated in **Figure 3**, shows a steady increase in performance from 2020 to 2024. Perceived effectiveness increased from 40% in 2020 to 70% in 2024, showing improvements in these tools’ ability to detect AI-generated content. Despite this upward trend, effectiveness remains below 100%, suggesting that current AI detection tools must fully address the complexities of detecting advanced AI outputs and that advancements are needed to keep pace with evolving AI technologies.

As detailed in **Table 8**, **Table 8** compares advanced AI detection tools, highlighting their efficacy and limitations in identifying AI-generated content. These tools are being utilized more frequently in educational settings to address the growing challenge of maintaining academic integrity in the face of AI-assisted cheating. Each tool has different capabilities, and while they provide valuable support, they also have limitations that require ongoing refinement as AI technologies evolve.

The findings in this theme emphasize the growing importance of addressing academic integrity issues raised by AI tools such as ChatGPT. While detection technologies have advanced, there is room for

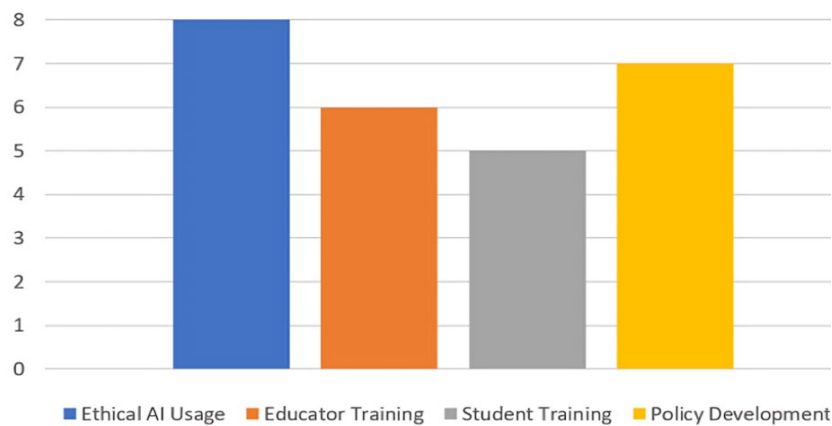


Figure 4. Frequency of comprehensive policy and training for AI usage mentioned, based on the findings of the 76 articles systematically reviewed in this study

improvement to ensure reliable identification of AI-generated content. It implies that detection tools alone may not be sufficient without complementary strategies such as innovative assessment redesign.

Comprehensive Policy and Training

Though AI detection tools are essential in maintaining academic integrity, their effectiveness relies on their integration into a larger framework of clear policies and robust training programs. Similarly, as exam design evolves to meet the challenges posed by AI, innovative assessment strategies can supplement these efforts by introducing more dynamic and personalized methods of assessing student learning. Adapting educational assessments to account for the increasing impact of AI tools such as ChatGPT requires enforcing innovative assessment strategies. As such, various studies emphasize the urgent need for HEIs to implement comprehensive policies and training programs in response to the rise of AI tools like ChatGPT. These policies guide students and educators in AI's ethical and practical use while maintaining academic integrity. Researchers suggest that HEIs must provide training programs for faculty and students to ensure that AI is used responsibly and that formal policies are enforced through standard guidelines and supervision (Gamage et al., 2023; Rudolph et al., 2024).

In addition, this study identified four major components that are essential for responsibly integrating AI into education. These components ensure that both educators and students have the knowledge and guidelines for using AI ethically, maintaining academic integrity, and adapting assessments to the challenges posed by AI tools such as ChatGPT. To ensure a smooth integration of responsible AI in HEI, researchers suggest the following measures on each component:

1. **Ethical AI usage:** Creating guidelines for ethical AI to promote academic integrity and ensure that AI tools such as ChatGPT improve learning rather than facilitate cheating. It also addresses issues of data privacy and fairness (Dai et al., 2023; Güner et al., 2024).
2. **Educator training:** Teaching educators how to integrate AI into their teaching and assessments to prepare them to effectively manage AI-related challenges, design robust assessments, and use AI to improve student engagement and learning outcomes (Gamage et al., 2023; Moorhouse et al., 2023).
3. **Student training:** Ensuring that students understand how to use AI tools ethically and responsibly to contribute to the development of an academic honesty culture and reduce the likelihood of AI misuse in academic work (Gonsalves, 2023; Susnjak & McIntosh, 2024).
4. **Policy development:** Developing clear institutional policies that address AI integration in education to provide a framework for upholding academic integrity standards, guiding students and faculty in the ethical use of AI, and ensuring adherence to academic guidelines (Chaudhry et al., 2023; Rudolph et al., 2024).

Consequently, **Figure 4** shows ethical AI usage as the most frequently mentioned component, which strongly emphasizes academic integrity and the responsible use of AI in education. Policy development is also highly relevant, underlining the importance of clear institutional guidelines. Educator training follows closely,

Table 9. Comprehensive policy framework for managing AI usage in education

Policy framework element	Description	Purpose	Considerations
Governance	Establish clear institutional rules and guidelines for the use of AI in education.	To ensure transparency, fairness, and consistency in how AI tools are used by both students and educators.	Regular policy reviews and updates to keep pace with AI advancements. Clear communication with all stakeholders (Farooqui et al., 2024).
Ethical guidelines	Develop a code of conduct outlining ethical AI usage. Include both acceptable and unacceptable AI use cases in academic work.	To promote responsible AI use, avoid academic dishonesty, and ensure students understand the boundaries of AI usage.	Collaboration with academic departments to align ethical AI usage with disciplinary standards (Gonsalves, 2023).
AI detection tools	Implement and regularly update AI detection software to identify AI-generated content in assignments and exams.	To help maintain academic integrity and detect potential misuse of AI tools in student submissions.	Integration into existing LMS. Training for educators on the limitations and appropriate use of detection tools (Nikolic et al., 2023).
Assessment design	Redesign assessments to prioritize higher-order cognitive skills, such as critical thinking, problem-solving, and creativity, which are more difficult for AI to replicate.	To reduce reliance on fact-based recall assessments that are easily manipulated by AI tools.	Use of diversified assessment methods, such as oral exams, project-based assessments, and reflective tasks (Klyshbekova et al., 2024).
Student and educator training	Provide comprehensive training programs on AI literacy, including how to use AI tools responsibly and the consequences of misuse.	To ensure both students and educators are knowledgeable about AI tools and the ethical considerations surrounding their use.	Ongoing workshops, webinars, and resources that evolve with AI technology. Incorporate training into student orientation and faculty development programs (Chinengundu, 2023).
Monitoring and enforcement	Establish a monitoring system to ensure compliance with AI usage policies and enforce disciplinary actions for violations.	To ensure that policies are adhered to and to maintain the credibility of academic programs.	Develop clear consequences for policy breaches and ensure transparent enforcement procedures. Use AI tools to monitor compliance in academic work (Dai et al., 2023).

reflecting the significance of providing educators with the skills they need to manage AI-related challenges, while student training is slightly lower, emphasizing the need to focus more on preparing students for ethical AI use.

Furthermore, Rudolph et al. (2024) support these suggested measures by highlighting the importance of institutions developing clear policies for AI use in education, including rules for when and how AI tools such as ChatGPT can be used. Similarly, Newton and Xiromeriti (2024) proposed evidence-based policies to protect the integrity of assessments and exams, ensuring that AI-generated content does not violate academic standards. Eventually, **Table 9** presents a comprehensive policy framework for managing the use of AI in educational settings. It concentrates on critical elements such as governance, ethical guidelines, AI detection tools, and assessment redesign, all of which aim to promote responsible AI usage while preserving academic integrity. This framework, which integrates clear guidelines, regular training, and innovative assessment strategies, provides a structured approach for HEIs to adapt to the changing landscape of AI technologies in academia. The findings in the last theme assert the critical importance of clear guidelines and robust training programs to ensure the responsible use of AI in educational settings. These policies and educator and student training are essential for upholding academic integrity and effectively incorporating AI technologies such as ChatGPT into the learning environment.

Overall, these findings highlight the potential and limitations of current strategies for maintaining academic integrity in the face of AI advancements. The following discussion will delve deeper into these findings, linking them to existing literature and identifying practical implications for educators and institutions.

DISCUSSION

Building on the findings presented above, this section will examine the implications of the identified strategies, beginning with effective exam design. These strategies, while promising, require further improvement to ensure that assessments remain strong in the face of evolving AI capabilities.

This study examined the implications of incorporating AI technologies, specifically ChatGPT, into educational settings. It investigated the challenges that these tools pose to traditional assessment methods and identified strategies, such as innovative exam design, AI detection tools, and comprehensive policy frameworks, that can help maintain academic integrity. The following themes provide a detailed analysis of how these approaches can improve the learning process while addressing ethical issues.

A primordial theme covered in this study is effective exam design. Gonsalves (2023) asserts the importance of designing assessments that can challenge AI systems and focus on assessing higher-order cognitive skills rather than just factual recall. Effective exam design requires assessments that can withstand AI tools like ChatGPT. Traditional exams based on rote memorization or simple factual recall are becoming increasingly ineffective (Kurtz et al., 2024). Instead, open-ended questions, oral exams, and project-based tasks requiring critical thinking and personal insight are proposed as more suitable alternatives. These formats make it more challenging for AI-generated content to succeed, ensuring students show genuine comprehension and creativity.

Similarly, Susnjak and McIntosh (2024) emphasize redesigning traditional exams to resist AI tools such as ChatGPT. Integrating complex multimodal questions with text and visuals bolsters assessments and decreases vulnerability to AI-generated content. Furthermore, Newton and Xiromeriti (2024) assert the importance of effective exam design, stating that MCQs, mainly in university assessments such as professional licensing exams, are crucial when assessing higher-order learning. When adequately structured, these MCQs cover a wide range of topics and can deliver more comprehensive assessments of student comprehension.

Another theme highlighted in the study is innovative assessment strategies. These strategies stimulate critical thinking, creativity, real-world problem-solving, and personalized learning, ultimately nurturing authentic student engagement and minimizing the risk of AI misuse, thus upholding academic integrity. Project-based learning, group discussions, and peer reviews are a few of its forms and serve as critical methods for combating the misuse of AI in education. It promotes active student participation and problem-solving, fostering higher-order cognitive abilities that AI tools cannot easily replicate. Changing away from traditional exam formats allows educators to ensure students engage more deeply with the material while promoting academic integrity.

Several authors discuss how innovative assessment strategies can promote critical thinking, creativity, and real-world problem-solving while mitigating the risk of AI misuse. For example, Gamage et al. (2023) highlight the potential of transitioning from traditional assessment methods to more interactive and applied assessments that require higher-order thinking and creativity, making them less susceptible to AI manipulation. They emphasize that such strategies can encourage students to engage more authentically while maintaining academic integrity. Likewise, Chaudhry et al. (2023) advocate for innovative assessment strategies encouraging group discussions, collaborative projects, and real-world problem-solving tasks. These methods stimulate critical thinking and creativity and make it more challenging for students to misuse AI tools like ChatGPT. Such approaches promote authentic student engagement and help maintain academic integrity.

Advanced AI detection tools are another theme that plays a pivotal role in the results of the study. As AI tools like ChatGPT become more sophisticated, these detection methods help educators avoid potential misuse, preserving the credibility and fairness of academic evaluations (Moorhouse et al., 2023). It is also critical to detect AI-generated content in scholarly work and maintain academic integrity as AI tools such as ChatGPT become more widely used. These detection methods include text analysis software, AI-enhanced plagiarism detection, and linguistic fingerprinting, which examines writing patterns to identify inconsistencies or AI involvement. The studies of Gamage et al. (2023) and Susnjak and McIntosh (2024) emphasize the need for sophisticated algorithms and AI-driven detection tools to detect AI-assisted cheating and ensure fair assessments effectively. In addition, Rudolph et al. (2024) and Richards et al. (2024) highlight tools like AI-enhanced plagiarism detection and advanced algorithms for their ability to detect subtle patterns in AI-

generated text that traditional methods may miss. These authors argue that as AI technology advances, detection tools must become more sophisticated to keep up with AI's growing capabilities in academic settings.

Moreover, **Table 8** compares advanced AI detection tools, highlighting their potential and limitations in addressing the growing issue of AI-generated academic misconduct. While tools like GPTZero, Turnitin's AI detection feature, and DetectGPT can help identify AI-generated content, their effectiveness is limited by challenges like detecting paraphrased or edited AI responses and the rapid evolution of AI models. These limitations highlight the need for a multifaceted approach combining detection tools with novel assessment designs prioritizing higher-order thinking and creativity, which AI systems struggle to replicate. Continuous improvement of these tools, combined with clear institutional policies and educator training, will be critical in maintaining academic integrity as AI technology advances.

Data from the study reinforces the last theme, which concerns comprehensive AI policy and training. Educators and institutions must implement comprehensive policies and training to adequately integrate AI technologies such as ChatGPT into education. The theme highlights the importance of clearly defined guidelines for the ethical use of AI and structured training programs for educators and students. Such policies are critical for upholding academic standards and ensuring AI tools enhance rather than sabotage learning. Dwivedi et al. (2021) and Perkins (2023) stress the importance of establishing clear guidelines on the ethical use of AI, alongside implementing training programs to equip educators and students with the knowledge to use these tools responsibly. However, many existing policies are outdated and lack sufficient clarity on the unauthorized use of AI technologies. With this, De Maio (2024) and Bobula (2024) argue that policies must undergo regular updates to keep pace with the rapidly evolving challenges AI presents.

In light of the need for updated policies, **Table 9** outlines a comprehensive policy framework for integrating AI into educational settings while preserving academic integrity. Each element, from governance to assessment design, highlights the need for transparent institutional policies, ethical guidelines, and AI detection tools to mitigate the misuse of AI tools like ChatGPT. The framework also highlights the importance of continuous educator and student training to ensure responsible AI usage. While the policy elements provide a robust foundation, the success of such a framework will rely on its adaptability to the rapid evolution of AI technologies and the effectiveness of enforcement mechanisms. This holistic approach guarantees that AI tools are harnessed for positive learning outcomes while minimizing risks to academic integrity.

To put it concisely, the four themes—effective exam design, innovative assessment strategies, advanced AI detection tools, and comprehensive policy and training—are interconnected in addressing the challenges posed by AI tools such as ChatGPT in education. Effective exam design and innovative assessments help turn the emphasis away from easily replicable tasks toward those demanding deeper understanding and creativity, lowering the risk of AI exploitation. At the same time, advanced detection tools help identify AI-generated content, and comprehensive policies and training ensure that educators and students have the knowledge and guidelines they need to use AI ethically. Together, these approaches reinforce the integrity and fairness of assessments in an AI-driven environment.

The results indicate clear pathways for implementing higher education assessment strategies. Faculty development programs are vital to equip educators with skills in designing AI-resistant assessments. Educators can effectively scale oral exams by leveraging video conferencing tools such as Zoom or Microsoft Teams, facilitating accessibility and real-time interaction. Similarly, project-based learning and multimodal assessments can incorporate collaborative tools like Google Workspace or Microsoft Teams to enable student collaboration, even in remote or hybrid learning environments. These approaches align with institutional goals to enhance critical thinking, creativity, and real-world problem-solving, ensuring assessments reflect genuine student abilities. By integrating these practical strategies, institutions can address AI's challenges while promoting authentic and ethical learning experiences.

Integration of Recommendations Across Educational Settings

Educators can adapt the proposed recommendations to traditional classroom settings by incorporating oral exams and encouraging students to express their critical thinking and reasoning skills during face-to-face interactions. For example, science students could explain experimental methods, while literature students

might critique the themes of a novel. Faculty members can use standardized rubrics for evaluation and participate in training workshops to develop practical oral exam questions to assure consistency.

In online learning environments, educators can leverage platforms such as Zoom or Microsoft Teams to conduct real-time oral assessments, where students present their ideas and respond to follow-up questions in virtual breakout rooms. Project-based learning is also highly flexible online, permitting students to collaborate asynchronously or synchronously using tools like Google Workspace to integrate text, visuals, and multimedia in their submissions. To encourage these approaches, institutions can develop guidelines for remote assessments, focusing on technical requirements such as internet stability, time slot scheduling, and utilizing recording features for post-assessment reviews.

Similarly, hybrid educational models offer another avenue for executing these strategies, combining in-person and virtual formats to create flexible and engaging assessment designs. Students can present findings in live sessions while submitting complementary reports or videos through learning management systems (LMS). Group projects in hybrid settings could involve some team members collaborating in person while others participate virtually, ensuring inclusivity and accessibility. Institutions can support these hybrid assessments by supplying flexible submission options and streamlining processes for faculty and students.

The recommendations also have discipline-specific applications. In STEM courses, oral exams involve explaining design processes or troubleshooting code errors, while biology students could discuss experimental results. In the humanities course, educators can use oral discussions or creative multimodal presentations to help students analyze historical events or interpret philosophical arguments. In business and law, project-based learning could feature simulations or case studies where students examine real-world business challenges or argue legal cases. These discipline-specific implementations can align assessment designs with each field's unique competencies and learning outcomes, assuring relevance and effectiveness.

To sustain the widespread implementation of these methods, institutions must provide the required infrastructure and tools, including access to collaborative software and video conferencing platforms. Training programs and dedicated support teams can aid educators in transitioning to innovative assessments. Institutions should also consider gradually adopting these methods by piloting them in smaller courses before scaling up across larger programs. This phased approach allows institutions to refine their practices while addressing logistical challenges, confirming a seamless transition to these forward-thinking assessment strategies.

CONCLUSION

AI technologies, particularly ChatGPT, have transformed the educational landscape, presenting opportunities and challenges for HEIs. This study highlights the urgent need for HEIs to adapt their assessment frameworks to uphold academic integrity amidst these advancements. As outlined in the introduction, addressing academic integrity requires a multi-faceted approach, including rethinking exam design, implementing innovative assessment strategies, adopting advanced AI detection tools, and establishing comprehensive institutional policies and training.

In alignment with these objectives, the findings emphasize transitioning from traditional fact-based assessments to formats prioritizing higher-order cognitive skills, such as critical thinking, creativity, and real-world problem-solving. These approaches reduce vulnerability to AI misuse while promoting authentic student engagement. The integration of advanced AI detection tools, as discussed, enhances the ability to identify AI-generated content, thereby safeguarding the authenticity of academic submissions.

Furthermore, the study emphasizes the importance of ethical AI policies and training programs for educators and students. By fostering AI literacy and promoting responsible AI use, these initiatives create a culture of academic integrity that is critical in guiding the challenges posed by AI technologies.

While AI tools like ChatGPT present challenges, they also offer significant opportunities to enrich the educational experience through personalized learning and innovative assessment methods. Looking ahead, HEIs must remain proactive, continuously refining their strategies to keep pace with rapidly evolving AI technologies. By doing so, institutions can ensure that academic credentials remain credible and that assessments reflect students' authentic abilities in this AI-driven era.

Limitations

While this study provides an in-depth examination of the challenges and strategies for maintaining academic integrity in AI tools such as ChatGPT, several limitations should be noted. First, the study's scope is limited primarily to ChatGPT, with little coverage of other emerging AI models and technologies that may pose similar or new challenges in higher education. However, the findings of this study, despite their limitations, can still provide valuable insights and strategies for maintaining academic integrity in AI tools and platforms.

Second, the study is based primarily on existing literature, which means that real-world empirical data on the long-term effects of implementing the recommended exam design and AI detection strategies in diverse educational settings is urgently needed. This reliance on secondary sources restricts the ability to measure the practical impact of these approaches in live classroom environments, highlighting the importance of further research in this area.

Third, the study concentrates on HEIs and does not account for potential differences in how AI tools impact academic integrity in other educational contexts, such as K-12 education or vocational training. The unique challenges and solutions in these other sectors may require further exploration.

Finally, the rapid evolution of AI technology limits the study's long-term relevance. As AI tools advance, new forms of academic misconduct may emerge, necessitating the evolution of detection tools. This study provides a snapshot of the current landscape, but further research will be required to stay ahead of these technological developments.

Future Directions

Building on this study's findings, several avenues for future research and practice are essential to further address the challenges posed by AI technologies in education. First, future studies should look beyond ChatGPT to see how other emerging AI tools and language models affect academic integrity. As new AI technologies with varying capabilities emerge, there is a need for comprehensive research that investigates their specific challenges and opportunities in a variety of educational contexts.

Second, empirical research is needed to determine the effectiveness of the proposed exam redesigns and AI detection tools in real-world educational setups. Longitudinal studies that track the effects of innovative assessment strategies like project-based learning, oral exams, and real-time feedback mechanisms will provide helpful information about their impact on student learning, engagement, and academic integrity. Furthermore, field experiments to evaluate the accuracy and adaptability of AI detection software in detecting new types of AI-generated content are critical.

Third, future research should investigate the creation and integration of more sophisticated AI detection tools that can keep up with AI models' rapidly evolving capabilities. To ensure widespread adoption, these tools must be adaptable, detect nuanced AI-generated content, and seamlessly integrate with existing LMS. Fourth, the role of ethical frameworks in guiding the responsible use of AI in education demands more attention. Researchers should explore how HEIs can develop comprehensive, evidence-based policies that prevent AI misuse and promote AI technologies' ethical and constructive use. This includes investigating best practices for training educators and students in AI literacy and responsible use.

Finally, future research should look beyond higher education to see how AI affects academic integrity in other educational sectors like K-12, vocational training, and professional certification programs. Each sector may face unique challenges requiring customized solutions for assessment design, policy development, and AI tool integration.

Future research can address these issues and help us better understand how to maintain academic integrity in an AI-driven educational landscape while maximizing the benefits of these technologies.

Funding: The author received no financial support for the research and/or authorship of this article.

Ethics declaration: This study does not require any ethical approval.

Declaration of interest: The author declares no competing interest.

Data availability: Data generated or analyzed during this study are available from the author on request.

REFERENCES

- Ali, D., Fatemi, Y., Boskabadi, E., Nikfar, M., Ugwuoke, J., & Ali, H. (2024). ChatGPT in teaching and learning: A systematic review. *Education Sciences*, 14(6), Article 643. <https://doi.org/10.3390/educsci14060643>
- Bobula, M. (2024). Generative artificial intelligence (AI) in higher education: A comprehensive review of challenges, opportunities, and implications. *Journal of Learning Development in Higher Education*, (30). <https://doi.org/10.47408/jldhe.vi30.1137>
- Bukar, U. A., Sayeed, M. S., Razak, S. F. A., Yogarayan, S., & Amodu, O. A. (2024). An integrative decision-making framework to guide policies on regulating ChatGPT usage. *PeerJ Computer Science*, 10, Article e1845. <https://doi.org/10.7717/peerj-cs.1845>
- Chaudhry, I. S., Sarwary, S. A. M., El Refae, G. A., & Chabchoub, H. (2023). Time to revisit existing student's performance evaluation approach in higher education sector in a new era of ChatGPT–A case study. *Cogent Education*, 10(1), Article 2210461. <https://doi.org/10.1080/2331186X.2023.2210461>
- Chinengundu, T. (2024). Higher education assessment at crossroads on the use of ChatGPT. *Proceedings of the International Conference on Future of Education*, 6(1), 87–98.
- Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228–239. <https://doi.org/10.1080/14703297.2023.2190148>
- Dai, Y., Liu, A., & Lim, C. P. (2023). Reconceptualizing ChatGPT and generative AI as a student-driven innovation in higher education. *Procedia CIRP*, 119, 84–90. <https://doi.org/10.1016/j.procir.2023.05.002>
- De Maio, C. (2024). Institutional responses to ChatGPT: Analysing the academic integrity policies of four public and private institutions of higher education in Australia. *Journal of Academic Language and Learning*, 18(1), T1–T8.
- Dwivedi, Y. K., Hughes, D. L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., ..., & Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice, and policy. *International Journal of Information Management*, 57, Article 101994. <https://doi.org/10.1016/j.ijinfomgt.2019.08.002>
- Farooqui, M. O., Siddiquei, M. I., & Kathpal, S. (2024). Framing assessment questions in the age of artificial intelligence: Evidence from ChatGPT 3.5. *Emerging Science Journal*, 8(3), 948–956. <https://doi.org/10.28991/ESJ-2024-08-03-09>
- Floridi, L. (2023). AI as *agency without intelligence*: On ChatGPT, large language models, and other generative models. *Philosophy & Technology*, 36, Article 15. <https://doi.org/10.1007/s13347-023-00621-y>
- Gamage, K. A. A., Dehideniya, S. C. P., Xu, Z., & Tang, X. (2023). ChatGPT and higher education assessments: More opportunities than concerns? *Journal of Applied Learning and Teaching*, 6(2), 358–369. <https://doi.org/10.37074/jalt.2023.6.2.32>
- Gonsalves, C. (2023). On ChatGPT: What promise remains for multiple choice assessment? *Journal of Learning Development in Higher Education*, (27). <https://doi.org/10.47408/jldhe.vi27.1009>
- Güner, H., Er, E., Akçapınar, G., & Khalil, M. (2024). From chalkboards to AI-powered learning: Students' attitudes and perspectives on use of ChatGPT in educational settings. *Educational Technology & Society*, 27(2), 386–404.
- Halaweh, M. (2023). ChatGPT in education: Strategies for responsible implementation. *Contemporary Educational Technology*, 15(2), Article ep421. <https://doi.org/10.30935/cedtech/13036>
- Imran, M., & Almusharraf, N. (2023). Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature. *Contemporary Educational Technology*, 15(4), Article ep464. <https://doi.org/10.30935/cedtech/13605>
- Klyshbekova, M., & Abbott, P. (2024). ChatGPT and assessment in higher education: A magic wand or a disruptor? *Electronic Journal of e-Learning*, 22(2), 30–45. <https://doi.org/10.34190/ejel.21.5.3114>
- Kurtz, G., Amzalag, M., Shaked, N., Zaguri, Y., Gal, E., & Zailer, G. (2024). Strategies for integrating generative AI into higher education: Navigating challenges and leveraging opportunities. *Education Sciences*, 14(5), Article 503. <https://doi.org/10.3390/educsci14050503>
- Lancaster, T. (2024). A large language model supported synthesis of contemporary academic integrity research trends. *ArXiv*. <https://doi.org/10.48550/arXiv.2401.03481>

- Lo, C. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, 13(4), Article 410. <https://doi.org/10.3390/educsci13040410>
- Mapletoft, N., Price, A., Smith, K., Mapletoft, O., & Elliott, M. (2024). An attempt to cheat using GPT-4: Findings, discussion and recommendations for academic staff and students. *Enhancing Teaching and Learning in Higher Education*, 1, 52–73. <https://doi.org/10.62512/etlhe.11>
- McIlwraith, T., Finnis, E., & Jones, S. (2023). Artificial intelligence, academic misconduct, and the borg: Why GPT-3 text generation in the higher education classroom is becoming scary. *Anthropologica*, 65(1), 1–27. <https://doi.org/10.18357/anthropologica65120232166>
- Meyer, J. G., Urbanowicz, R. J., & Martin, P. C. N. (2023). ChatGPT and large language models in academia: Opportunities and challenges. *BioData Mining*, 16, Article 20. <https://doi.org/10.1186/s13040-023-00339-9>
- Montenegro-Rueda, M., Fernández-Cerero, J., Fernández-Batanero, J. M., & López-Meneses, E. (2023). Impact of the implementation of ChatGPT in education: A systematic review. *Computers*, 12(8), Article 153. <https://doi.org/10.3390/computers12080153>
- Moorhouse, B. L., Yeo, M. A., & Wan, Y. (2023). Generative AI tools and assessment: Guidelines of the world's top-ranking universities. *Computers and Education Open*, 5, Article 100151. <https://doi.org/10.1016/j.caeo.2023.100151>
- Newton, P., & Xiromeriti, M. (2024). ChatGPT performance on multiple choice question examinations in higher education. A pragmatic scoping review. *Assessment & Evaluation in Higher Education*, 49(6), 781–798. <https://doi.org/10.1080/02602938.2023.2299059>
- Nikolic, S., Daniel, S., Haque, R., Belkina, M., Hassan, G. M., Grundy, S., Lyden, S., Neal, P., & Sandison, C. (2023). ChatGPT versus engineering education assessment: A multidisciplinary and multi-institutional benchmarking and analysis of this generative artificial intelligence tool to investigate assessment integrity. *European Journal of Engineering Education*, 48(4), 559–614. <https://doi.org/10.1080/03043797.2023.2213169>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ..., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *The BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Perkins, M. (2023). Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching and Learning Practice*, 20(2). <https://doi.org/10.53761/1.20.02.07>
- Plata, S., De Guzman, M., & Quesada, A. (2023). Emerging research and policy themes on academic integrity in the age of Chat GPT and generative AI. *Asian Journal of University Education*, 19(4), 743–758. <https://doi.org/10.24191/ajue.v19i4.24697>
- Rahimi, R., Jones, J., & Bailey, C. (2024). Exploring contract cheating in further education: Student engagement and academic integrity challenges. *Ethics and Education*, 19(1), 38–58. <https://doi.org/10.1080/17449642.2023.2299193>
- Richards, M., Waugh, K., Slaymaker, M., Petre, M., Woodthorpe, J., & Gooch, D. (2024). Bob or bot: Exploring ChatGPT's answers to university computer science assessment. *ACM Transactions on Computing Education*, 24(1), Article 5. <https://doi.org/10.1145/3633287>
- Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching*, 6(1), 342–363. <https://doi.org/10.37074/jalt.2023.6.1.9>
- Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning & Teaching*, 6(1). <https://doi.org/10.37074/jalt.2023.6.1.17>
- Susnjak, T., & McIntosh, T. R. (2024). ChatGPT: The end of online exam integrity? *Education Sciences*, 14(6), Article 656. <https://doi.org/10.3390/educsci14060656>
- Talan, T., & Kalinkara, Y. (2023). The role of artificial intelligence in higher education: ChatGPT assessment for anatomy course. *International Journal of Management Information Systems and Computer Science*, 7(1), 33–40. <https://doi.org/10.33461/uybisbbd.1244777>

Vargas-Murillo, A., Pari-Bedoya, I., & Guevara-Soto, F. (2023). Challenges and opportunities of AI-assisted learning: A systematic literature review on the impact of ChatGPT usage in higher education. *International Journal of Learning, Teaching and Educational Research*, 22(7), 122-135. <https://doi.org/10.26803/ijlter.22.7.7>

