INTELLIGENT TECHNOLOGIES IN EDUCATION

# Artificial intelligence in higher education learning: transferable skills and academic integrity

Toh Yen Pang[a], Alex Kootsookos[a], Chi-Tsun Cheng[a]

[a] *School of Engineering, STEM College, RMIT University, Melbourne, VIC 3000, Australia*

## Abstract

The advancement of Generative Artificial Intelligence (AI) chatbots, such as ChatGPT, presents significant and transformative challenges in higher education teaching and learning, such as assessment and evaluation practices. While this is acknowledged, there has been very little research into what this might look like in daily practice in higher education. This study explored these challenges in one area of higher education practice: developing students' transferable skills, including writing, critical thinking, and information literacy among undergraduate engineering students at RMIT University, Melbourne, Australia. Using a cohort comparison design, this study evaluated the impact of ChatGPT on students' attainment of transferable skills. The effectiveness of AI tools in enhancing educational outcomes was assessed with a standardised assessment framework used by independent assessors to grade students' reports. The results, analysed using the Mann-Whitney *U* test and the intraclass correlation coefficient, revealed significant improvements in critical thinking and information literacy among those students who used ChatGPT. The study also explored the ethical implications of using AI in educational settings and highlighted the need for rigorous academic standards and the implementation of measures to ensure the responsible use of AI technologies. While the preliminary findings suggest that AI tools, particularly ChatGPT in this study, can positively impact certain students' skills, more detailed and controlled studies are necessary to validate these results and explore further the mechanisms through which AI tools influence learning and skill development.

**Editors**

Section: Educational Technology

Co-Editor-in-Chief: Dr Marios Kremantzis

Associate Editor: Dr Xue Zhou

**Publication**

Received: 8 October 2024
Revised: 16 January 2025
Accepted: 3 March 2025
Online: 4 March 2025

## Keywords

Generative Artificial Intelligence (GenAI), Pedagogical Practices, Learning and Teaching, Higher Education, Engineering

## Citation

Pang, T. Y., Kootsookos, A., & Cheng, C. T. (2025). Artificial Intelligence in Higher Education Learning: Transferable Skills and Academic Integrity. Intelligent Technologies in Education, Advanced Online Publication.

# Introduction

Generative Artificial Intelligence (hereafter, AI) chatbots, notably ChatGPT—an AI chatbot created by OpenAI, Inc., San Francisco, California, USA—have significantly impacted assessment and evaluation practices in higher education, revolutionising the field and presenting both opportunities and challenges (Lytras et al., 2024; Peres et al., 2023). This impact requires a thorough examination of AI's transformative potential, as well as its ethical, practical, and educational challenges (Nguyen et al., 2024). One example, suggested by the Bioethics Centre at the University of Otago (Zohny et al., 2023), is AI technologies' capacity to generate human-like text that could potentially allow a broader range of authors to publish their ideas in humanities journals. This potential is especially critical for non-native English speakers, thus facilitating their contribution to scholarly discourse. Moreover, various educational disciplines, from science and medicine to economics and finance, are integrating AI into instructional applications (Batista et al., 2024; Lee et al., 2024). With its proficiency in processing and generating natural language, AI enhances learning and teaching by facilitating targeted questions and customising educational materials (Cooper, 2023; Ellikkal & Rajamohan, 2024; Khan et al., 2023; Lim et al., 2023; Peres et al., 2023; Rahman & Watanobe, 2023).

Nguyen et al. (2024) has recently explored both the opportunities and challenges presented by AI technology in higher education. They provided insights into how AI technologies can enhance student engagement while also raising ethical and pedagogical concerns. These findings align with recent discussions on AI's role in assessment and feedback, particularly in higher education settings (Roe et al., 2024). Research on AI-driven feedback suggests that while AI can provide rapid and automated responses, its effectiveness is significantly enhanced when combined with instructor feedback. Roe et al. (2024) found that both students and academic staff were sceptical of AI-generated feedback when used autonomously, but were more open to AI-assisted feedback in combination with human input. This suggests that

AI’s role in education should be viewed as augmenting, rather than replacing, traditional feedback mechanisms.

However, the fast evolving and advancement of AI technologies has prompted many to raise critical questions about the future of learning and teaching with AI technologies (Baidoo-Anu et al., 2024; Essien et al., 2024; Lee et al., 2024; Lytras et al., 2024). There are concerns about AI’s capabilities to generate rapid human-like answers, especially in tasks such as essay writing and solving mathematical and/or programming problems, where AI might lead to academic dishonesty among learners (Lee et al., 2024). Students might rely on it to complete assignments and exams without genuine effort or understanding, which could compromise the learning process and research authenticity. Similarly, educators might use it to generate lesson plans and answers to student queries rather than as an aid, leading to a potential degradation in the quality of teaching practice (Buruk, 2023; Cooper, 2023; Cotton et al., 2023; Khan et al., 2023; Lim et al., 2023; Megahed et al., 2023; Rahman & Watanobe, 2023).

Consequently, educators and policymakers grapple with how best to leverage these AI capabilities for assessment and evaluation purposes and deal with the potential downsides. As noted in many recent studies of university students' and staff’s perceptions of AI technologies, the common consensus on the deployment of AI in higher education is that it could provide personalised learning support, writing assistance, and enhance research capabilities (Chan, 2023; Chan & Hu, 2023; Cooper, 2023; Dai, Liu, et al., 2023; Ellikkal & Rajamohan, 2024; Lim et al., 2023). However, concerns with reliability, privacy, and ethical issues have also emerged, influencing students' perceptions of AI in their learning outcomes and their engagement (Lodge, 2024; Megahed et al., 2023; Qadir, 2023; Wang et al., 2023).

As such, these issues emphasise the need for responsible use, critical thinking, and understanding of AI’s limitations in an assessment context. The advent of AI in assessment practices calls for a nuanced understanding and a holistic approach to its integration within

academic settings to ensure effective learning outcomes while addressing ethical and practical issues (Alin et al., 2023; Chan & Hu, 2023; Cotton et al., 2023; Ellikkal & Rajamohan, 2024).

### Objective and research questions

One of the key challenges educators face is using AI-generated information effectively to enhance student learning outcomes whilst upholding ethical conduct and avoiding plagiarism. This project aimed to investigate the effectiveness and adaptability of AI in enhancing transferable skills among learners and the pedagogical strategies adopted by educators in higher education to guide learners to avoid plagiarism and to understand the best practices for incorporating AI in their learning journeys. The main research questions that this project aimed to address were:

RQ1: *How can ChatGPT or similar AI-based chat prompts effectively enhance students' transferable skill development?*

RQ2: *What are the best practices for integrating ChatGPT into pedagogical practices to promote academic integrity and transferable skills?*

This study examined transferable skills, such as writing, critical thinking, information literacy, academic integrity and creative thinking, based on these contexts.

## Literature

### Generative Artificial Intelligence in education

Anastasia et al. (2023) examined the use and implications of Contextual Large Language Models (C-LLMs) such as GPT-3 in education. The authors presented two perspectives on using AI in education. First, AI models in their current form are potentially harmful to the epistemological foundations of modern education because they cannot filter unreliable facts or sources and lack ethical and critical thinking abilities. Secondly, when used appropriately, AI models can provide helpful feedback to students rapidly, helping them to learn. Despite its weaknesses, AI can be a beneficial tool if used properly.

In a similar vein, Kamalov et al. (2023) highlighted that AI tools offer immediate feedback and support, which is essential for developing communication skills and encouraging a proactive approach to learning. They also found that AI-based chat prompts can tailor educational content to individual student needs and consequently help cultivate critical thinking and problem-solving skills, which are essential in the workforce.

A recent study by Nyale et al. (2024) considered using AI-based chat prompts to facilitate personalised learning experiences for students in tertiary institutions. They aimed to improve students' transferable skills, such as critical thinking, adaptability and problem-solving. Through the study, the authors developed AI frameworks to align academic programs to cultivate transferable skills that prepare students for a dynamic digital economy and enhance their employability.

Farazouli et al. (2023) revealed important implications of AI chatbots, specifically ChatGPT, on university teachers' assessment practices and perceptions. Experiments showed that ChatGPT-generated answers achieved variable passing rates, indicating diverse capabilities in responding to examination prompts. The assessment outcomes showed that human assessors tend to identify fewer issues in chatbot-generated answers, while most critical comments were given to human-written submissions.

The advent of AI chatbots has also triggered policy changes and the creation of guidelines in higher education institutions, influencing institutional practices and staff development. Despite potential benefits, the introduction of AI chatbots, e.g. ChatGPT, poses challenges that need to be addressed before educators can fully embrace those technologies and incorporate them seamlessly into higher education.

Sullivan et al. (2023) reviewed public discussions and university responses in higher education to AI, specifically OpenAI's ChatGPT. They investigated academic integrity concerns, the limitations of AI tools and potential opportunities for student learning. Their analysis was based on a content analysis of 100 news articles from Australia, New Zealand, the US and the UK.

Their findings suggested that some academics believed that the introduction of ChatGPT has raised substantial academic integrity concerns because it could help students uncritically write assignments. However, other academics believe that it could help students learn better and that they should teach students to use AI tools properly and ethically.

There has been little discussion about the potential benefits of AI tools to improve participation and success rates among students from disadvantaged backgrounds. Moreover, the literature often neglects students' perspectives. Despite considerable media attention, academic literature has given little focus to the impact of sophisticated AI tools in higher education. This gap highlights the need for further research on how AI tools are transforming education and how such technologies can address the diverse needs of students.

**Perspectives of educators and learners on AI in education**

The adoption and perception of AI technologies in higher education reveal significant regional and cultural contexts, with educators and students often perceiving their benefits and challenges differently. Understanding their perspectives provides insights into gaps in their perceptions and highlights how these differences shape the integration and utilisation of AI tools in educational contexts.

***Student Perspectives***

Studies on students' perspectives emphasise regional variations in awareness, usage and concerns regarding AI technologies. Students generally have positive perceptions of using AI language tools, and many claim that AI makes them effective learners (Kurtz et al., 2024). Johnston et al. (2024) observed that UK students were highly aware of AI-assisted tools, with over 93% of surveyed students recognising and supporting using such tools (e.g. Grammarly and ChatGPT). While AI tools were widely accepted, students opposed using ChatGPT to produce entire essays. These findings emphasised the need for university-wide policies to guide ethical use. Similarly, Baidoo-Anu et al. (2024) reported that Ghanaian students were aware of ChatGPT but primarily used it for assignments rather than collaborative projects. They found that limited training on safe and effective usage was a barrier. The findings

reported that students are concerned about academic misconduct, originality and data security.

In contrast, Kim et al. (2024) found that Chinese students in China viewed AI as a multitasking assistant, virtual tutor and digital peer that improved their writing performance and engagement. However, they were cautious about over-reliance. Stone (2024) extended this discussion to the United States, finding mixed student responses. While many were excited about the potential of AI, concerns about fairness, accessibility and its ambiguous role in assessments were frequently noted. Gender disparities in adoption also emerged, with men reporting higher usage rates than women, indicating sociocultural factors influencing adoption.

***Educator Perspectives***

Educators' perspectives on the adoption and impact of AI tools, such as ChatGPT, are shaped by the tools' growing influence on assessment practices and the broader educational landscape (Batista et al., 2024; Lee et al., 2024). There is also a lack of unified consensus among educators regarding AI in higher education, along with ambiguity about best practices related to recent technological advancements (Lee et al., 2024). For example, Kizilcec et al. (2024) conducted a cross-continental study (Australia, Cyprus and the United States) and found that educators viewed AI as a disruptor to traditional assessment practices. While educators strongly supported redesigning assessments to incorporate AI, they expressed concerns about its potential to undermine academic authenticity and creativity. These findings emphasise the need for collaborative efforts between educators and students to reform assessment practices. Similarly, Pang et al. (2024) examined non-English-speaking background educators in Australia and found that linguistic proficiency and cultural background significantly shaped attitudes toward AI. While educators acknowledged AI's potential to facilitate the prescription of feedback to students, they also emphasised the need for human oversight, staff training and alignment with individual teaching styles.

In the Middle East, Khlaif et al. (2024) explored the perspectives of early-adopter instructors using AI and found that institutional policies and cultural norms significantly influenced usage.

Educators developed systematic approaches to integrate AI into assessments, focusing on reducing workload and improving engagement. However, concerns about ethical implications and student over-reliance on AI remained challenges, highlighting the need for structured support for educators and students navigating this technology.

Kaplan-Rakowski et al. (2023) investigated US educators and noted that frequent AI users had more positive attitudes toward its integration. Educators viewed AI as a tool for fostering critical thinking and enhancing professional development. However, they also expressed concerns about equity, ethical misuse, and the risk of de-skilling students in essential areas such as writing and critical analysis.

There are notable differences between educators' and students' perspectives on AI in education. Educators tend to view AI as a tool for pedagogical enhancement and assessment. They are interested in integrating AI to promote critical thinking and authentic learning experiences (Kaplan-Rakowski et al., 2023; Kizilcec et al., 2024). Students, on the other hand, demonstrate a mix of enthusiasm and apprehension towards AI. While students recognise the potential benefits of improving learning processes, they are concerned about ethical implications, dependency and the impact on their skills development (Barrett & Pack, 2023; Johnston et al., 2024; Kim et al., 2024).

**Impact of AI on transferable skills**

Despite the widespread application of AI in various sectors, its implementation in education is currently limited, with a predominant use of AI technologies supplementary to the learning process, such as machine translation and grammar-checking tools (Cope & Kalantzis, 2023; Kaplan-Rakowski et al., 2023). However, the potential of AI to enhance educational practices is being increasingly explored, with the prospect of AI augmenting collaborative, immersive and exploratory learning experiences (Lane et al., 2016; Luckin, 2018; Markauskaite et al., 2022).

Numerous studies (Ellikkal & Rajamohan, 2024; Kim et al., 2024; Song & Song, 2023; Yilmaz

& Karaoglan Yilmaz, 2023) have demonstrated the potential of AI-driven educational tools in enhancing student engagement and personalized learning experiences, which are critical for developing deep analytical and problem-solving skills. These studies typically focused on AI's capability to provide personalised learning experiences and adaptive feedback, which are beneficial for skill acquisition.

While AI tools can provide immediate feedback and content suggestions (Dai, Lin, et al., 2023; Qadir, 2023; Zhao et al., 2023), they do not replace the need for critical thinking and problem-solving skills that are vital for genuine learning. Similarly, according to recent insights by Nikolic et al. (2024), AI-driven tools enabled students to meet basic educational requirements, but students required further exploration and analysis to achieve higher cognitive outcomes. However, little research has been done to explore how AI tools can be integrated with traditional teaching methods to enhance transferable skills.

Song and Song (2023) assessed the efficacy of ChatGPT in enhancing academic writing by Chinese students who took English as a Foreign Language (EFL). The findings suggested that integrating AI tools such as ChatGPT in language learning can significantly aid EFL students in developing writing proficiency. Additionally, these tools can help enhance broader skills such as critical thinking and information literacy, which are essential for academic success. Song and Song also suggested further research to explore how different types of AI tools might be optimised to support various aspects of learning and skills development, especially in settings that involve creative thinking tasks.

Transferable skills, often described as soft skills, are versatile abilities that enable individuals to perform effectively across various academic, professional and personal contexts. These skills are highly valued by employers for their role in promoting adaptability and lifelong learning (Huq & Gilbert, 2017; Rosenberg et al., 2012). In the context of our study, several authors identified transferable skills such as writing, critical thinking, information literacy (Burns et al., 2022; Welsh & Wright, 2010), academic integrity and creative thinking, based on

their broad applicability and critical role in both academic and professional environments.

### *Writing Skills*

Writing skills are widely recognised as transferable because they are crucial for effective communication. Organising ideas coherently, presenting arguments logically, and adhering to academic conventions are essential for success in various disciplines and professional contexts (Timmerman et al., 2011). AI tools have shown significant potential in developing students' writing skills, particularly in educational settings where English is a second language (Song & Song, 2023). Their findings revealed that AI tools enhance various aspects of writing, including grammar, coherence, organisation and vocabulary, which are critical for effective communication in academic and professional settings.

### *Critical Thinking skills*

Critical thinking involves analysing information objectively, evaluating evidence and constructing reasoned arguments. It is regarded as a core graduate outcome in higher education and is crucial for academic success and employability (Martín-Raugh et al., 2023; Paulsen, 2013). Recently, AI tools have sparked debate about their impact on critical thinking development. For instance, Essien et al. (2024) investigated the role of AI text generators in fostering critical thinking among UK postgraduate business students. The study revealed that AI tools were most effective at improving lower-order cognitive skills, such as knowledge recall and comprehension. However, their influence on higher-order skills, such as analysis, evaluation and creation, was limited.

### *Information Literacy*

The American Library Association defines information literacy as skills enabling individuals to recognise when information is needed and to locate, evaluate and use it effectively (Welsh & Wright, 2010). Higher education plays a critical role in developing students' information competencies. Burns et al. (2022) argued that systematic curriculum design and assessment are essential for embedding information literacy into degree programs and preparing students for real-world challenges. They also found that information literacy is often taught at a

foundational level but lacks reinforcement at advanced stages, limiting students' ability to transfer these skills effectively into digital dominant workforce.

***Creative Thinking skills***

Creative thinking involves generating original ideas, exploring multiple possibilities and connecting seemingly unrelated concepts. Harris et al. (2022) identified four interconnected elements crucial for process-oriented approaches to developing creative thinking: 1. Inquiring: identifying and clarifying questions and problems. 2. Generating ideas: exploring possibilities and actions to address challenges. 3. Reflecting on thinking processes: evaluating and improving one's creative approaches. 4. Analysing and synthesising: developing reasoning to justify and refine creative outputs. AI technologies offer both opportunities and challenges for fostering creative thinking in education. Anastasia et al. (2023) explored the potential of AI to enhance creativity by serving as a brainstorming partner, generating diverse ideas and simulating different perspectives. However, they also noted inherent limitations of AI, such as its reliance on pre-existing amount language, which can constrain novelty or context-specific creativity.

Understanding how AI tools influence these skills can provide valuable insights for educators and policymakers aiming to optimise learning outcomes and equip students for the rapidly evolving job market demands.

**Academic Integrity and Challenges of AI**

Academic integrity is a foundational transferable skill that upholds ethical principles, honesty and accountability in scholarly work. AI tools have introduced both opportunities to enhance learning and risks related to academic misconduct. Nikolic et al. (2024) conducted a multi-institutional study examining the performance of various Generative AI tools across engineering assessment types, such as quizzes, programming tasks and reflective writing. Their findings revealed that newer versions, like ChatGPT-4, performed exceptionally well across most assessment types, and there are heightened fears about their misuse in assessments. The traditional concept of academic integrity is evolving in response to the

capabilities of AI. Sullivan et al. (2023) suggested that universities must shift their focus from solely preventing misconduct to embracing AI's integration into teaching and learning. The process involves redefining academic integrity to encompass ethical AI use, transparency and the ability to critically evaluate AI-generated outputs.

A recent study by Darvishi et al. (2024) suggests that reliance on AI for content generation may inadvertently hinder the development of essential transferable skills if not integrated thoughtfully within educational frameworks. Research indicates that AI technologies, such as ChatGPT, can provide detailed, coherent and personalised feedback to students (Dai, Lin, et al., 2023; Qadir, 2023), support literacy development (Buruk, 2023; Li & Xu, 2023; Liu et al., 2023), and simulate student-teacher dialogues (Markel et al., 2023; Tack & Piech, 2022). However, the introduction of AI into educational practices also presents significant challenges. A primary concern is the potential for AI technologies to encourage academic dishonesty, with the capability of AI such as GPT-3 to produce human-like coherent text, thereby facilitating academic dishonesty by allowing individuals to generate complex, linguistically accurate content that is not their own (Cotton et al., 2023).

Moreover, AI's capacity to pass professional examinations, as seen in the US National Board of Medical Examiners test (Gilson et al., 2023; Kung et al., 2023), poses a significant challenge in determining the authenticity of students' work and knowledge. Traditional assessment methods may fall short, leading to the need for more stringent measures like invigilated assessments. However, such methods may only encourage rote learning and recall instead of fostering critical thinking skills.

As institutions adopt AI tools globally, especially ChatGPT, for various educational purposes, assessing the implications for ethical practice within academic settings becomes crucial. UNESCO (2023) emphasises the need for transparency, accountability and ethical considerations when integrating AI technologies into educational contexts. Hence, UNESCO's guidelines advocate equitable use to ensure that AI implementation does not exacerbate

educational inequalities or biases. The Australian Tertiary Education Quality and Standards Agency (TEQSA) has similarly emphasised the importance of clear policies, student training on ethical AI use, and rigorous assessment mechanisms to detect and deter academic misconduct (Lodge, 2024).

These concerns highlight the need for a careful, measured approach to incorporating AI into education. Our study aimed to shed light on effective strategies for achieving this balance, thereby preparing educators and learners for the transformative potential of AI in enhancing learning outcomes, personalising education and streamlining assessment strategies.

## Methodology

The data for this research were drawn from students studying different undergraduate engineering degrees at RMIT University in Melbourne, Australia, between 2022 and 2023.

### Research design

This study adopted a cohort comparison design to evaluate the effectiveness of ChatGPT in enhancing transferable skills among students. Students were categorised into two groups based on their exposure to ChatGPT:

1. Non-AI cohort: This group consisted of students from a cohort who completed their essays before the official launch of ChatGPT in November 2022 (Peres et al., 2023). These students did not have access to or were not instructed to use AI tools for their tasks.
2. AI cohort: This group included students introduced to ChatGPT as part of their coursework. They were provided essential guidance on using ChatGPT effectively for their essays.

### Participants

Table 1 presents the descriptive data for the participants of bachelor's in engineering honours degrees for the study population. More than 80% of the participants were male and largely were identified as Australian resident students.

**Table 1**

*Descriptive statistics for participants from the engineering degrees*

| Variable | Non-AI cohort (2022) | | AI cohort (2023) | |
|---|---|---|---|---|
| | N | % | N | % |
| Gender | | | | |
| Male | 189 | 89.2% | 191 | 82% |
| Female | 23 | 10.8% | 42 | 18% |
| Citizenships | | | | |
| Local students | 168 | 79.2% | 179 | 76.8% |
| International students | 44 | 20.8% | 54 | 23.2% |
| Engineering Honours Degrees | | | | |
| Aerospace | 12 | 6% | 111 | 48% |
| Advanced Manufacturing & Mechatronics | 87 | 41% | 66 | 28% |
| Automotive | 12 | 6% | 11 | 5% |
| Mechanical | 79 | 37% | 17 | 7% |
| Sustainable System | 8 | 4% | 19 | 8% |
| Others | 14 | 7% | 9 | 4% |

Both the non-AI and AI cohorts comprised students from diverse academic backgrounds. They were given the same task instructions to ensure consistency in expectations and to allow a fair comparison of outcomes. Students were asked to complete a Research Activity consisting of a written essay (600–650 words) on one of the following topics: 1. The process of refining iron ore to make steel. 2. The refinement of titanium ore to create a usable titanium alloy. 3. Metallic glasses or amorphous metals: their formation and applications. 4. Shape memory alloys and their use in engineering applications. 5. The concept of ‘superplasticity’ and its advantages in engineering.

Each student was allocated four weeks to complete the task, allowing sufficient time for research, drafting, and revisions. The same educator who taught both cohorts provided

identical standardised instructions to ensure that all students were given the same explanation of the essay requirements.

For the AI cohort, students were encouraged to utilise ChatGPT as part of their essay development process. Although direct monitoring of AI tool usage was not implemented, students were requested to include a brief summary of their search strategies and interactions with AI tools in their submissions. This strategy provided insight into how they approached and utilised ChatGPT while maintaining flexibility in its application.

**Ethics**

This study was approved by the RMIT University, Science, Technology, Engineering and Mathematics (STEM) College Human Ethics Advisory Network (CHEAN). Students were asked to consent to their reports being included anonymously in the dataset for analysis. The analysis for this study was conducted after the students' reports were marked, graded and returned to them. It was clear to the students that their participation or non-participation would not affect their grades. The course coordinator, i.e., the educator responsible for planning and coordinating course delivery, randomly selected ten essays from each to represent a diverse range of student performances. Anonymous identifiers (random codes) were assigned to the reports to ensure the anonymity of the participants was maintained.

**Selection of essays and Instruments for analysis**

A total of 20 essays were selected for in-depth analysis: 10 from the AI cohort and 10 from the non-AI cohort. Essays were chosen randomly from the two grading categories below to ensure a representative sample.

1. High Distinction (HD): Essays scoring 80–100%, demonstrating exceptional understanding, originality and critical analysis. These essays included well-formulated arguments, structured diagrams and relevant references.
2. Distinction (DI): Essays scoring 70–79%, showing a strong grasp of the subject matter, clear arguments, and some evidence of creative and solid work.

From each grading category, five essays were randomly chosen per cohort. This approach ensured a balanced representation of high-performing students across the two groups. To objectively assess and compare the transferable skills of students in both cohorts, an assessment framework was designed and developed based on previous literature to guide the evaluation process (Burns et al., 2022; Paulsen, 2013; Shively et al., 2018; Timmerman et al., 2011). A copy of the assessment framework can be found in Appendix A. The assessment framework was used to evaluate the key dimensions of students' transferable skills, such as writing, critical thinking, information literacy and creative thinking abilities.

Two independent accessors not involved in the course's teaching used this assessment framework to review and evaluate ten sample essays from each cohort. Assessors were trained in applying the framework to ensure consistency in evaluation standards and reliability of results. The assessors attended a workshop to familiarise themselves with the assessment framework. During this, they did three practical exercises to evaluate sample essays using the assessment framework. During the practical exercise, each assessor scored the sample essays, after which they compared their evaluations and noted the similarities and differences in their scoring. This process ensured that the assessors produced a high level of agreement in their scoring. Each assessor provided comments and supplied feedback on the student's essays.

**Data analysis**

In this study, the framework was used by two assessors working independently. To compare the ratings from the two assessors, an Interclass correlation coefficient (ICC) was used to determine inter-rater reliability (Alavi et al., 2022; Koo & Li, 2016). The initial assumption of equal variance (homogeneity) was tested using Fisher's F-test within the ICC calculations, allowing for a meaningful comparison of variability within and between groups, hence assessing consistency, or not, in assessors' ratings (McGraw & Wong, 1996).

To analyse the differences between non-AI and AI cohorts, a Mann-Whitney *U* test was conducted to compare the means of their transferable skills. The *p*-value < 0.05 was considered statistically significant for a confidence interval of 95%. Statistical analyses were performed using SPSS software version 29.0 (IBM® Corp. Armonk, NY, USA).

A qualitative content analysis approach was employed to analyse the independent assessors' comments and feedback on the student essays. Initially, the researchers familiarised themselves with the assessors' feedback to understand the central themes. A coding scheme was then established based on these themes and applied to the data. After that, the coded data were analysed to identify patterns and trends concerning the development of transferable skills among students in the context of the research questions (Mamabolo & Myres, 2019; Vogt et al., 2014). By adopting the qualitative content analysis approach, a nuanced understanding of the assessors' perspectives on student skills development provided valuable insights that can contribute to the enhancement of pedagogical practices in the future.

# Results

**Inter-rater reliability**

In analysing the inter-rater reliability for various skills, we rigorously analysed 20 exemplars of student essays, equally distributed between the non-AI and AI cohorts, which two assessors independently reviewed against a detailed set of criteria (Appendix A). Inter-rater reliability compared the level of agreement between different assessors. The findings, presented in Table 2, were interpreted against the guidelines within the established academic benchmarks (Alavi et al., 2022; Koo & Li, 2016).

**Table 2**

*Inter-rater reliability analysis of transferable skills assessment*

| Transferable skills | Intraclass correlation coefficient (ICC) | 95% Confidence Level | | F test with true value | | | |
|---|---|---|---|---|---|---|---|
| | | Lower Bound | Upper Bound | Value | *df 1* | *df 2* | Sig. |
| Writing skills | 0.533 | -0.72 | 0.808 | 2.345 | 19 | 19 | 0.035 |
| Critical thinking | 0.471 | -0.394 | 0.794 | 1.846 | 19 | 19 | 0.095 |
| Information literacy | 0.706 | 0.288 | 0.882 | 3.623 | 19 | 19 | 0.004 |
| Academic integrity | 0.676 | 0.207 | 0.870 | 3.152 | 19 | 19 | 0.008 |
| Creative thinking | 0.490 | -0.141 | 0.788 | 2.422 | 19 | 19 | 0.030 |
| Overall | 0.784 | 0.593 | 0.905 | 8.651 | 19 | 171 | <0.001 |

In the analysis of the student essays, ICC was employed as a statistical measure to evaluate the consistency or conformity of the quantifiable skills within the same class—in this case, the essays from the non-AI and AI cohorts. The ICC is a crucial indicator where a few assessors were involved because it helped to ascertain the degree to which these assessors provided consistent ratings to the same objects, i.e. the student essays.

It was observed that the ICC values for the various transferable skills ranged from moderate to good (i.e. below 0.5 indicated poor reliability, between 0.5 and 0.75 indicated moderate reliability, and between 0.75 and 0.9 indicated good reliability), suggesting different levels of agreement amongst the assessors. For example, writing skills showed an ICC value of 0.533, which was in the moderate range. This observation indicated that while there was some level of agreement among the assessors on the students' writing skills, there were also notable differences in how each assessor perceived and scored these skills. This level of ICC reflected a degree of subjectivity and individual variation in the assessment of writing skills, which is not uncommon in qualitative analyses where personal standards and interpretations may influence judgements.

Information literacy had a higher ICC value of 0.706, suggesting a strong consensus between assessors about the students' ability to analyse and synthesise information from various sources and support students' arguments with evidence. This high level of agreement was statistically supported by an F-test value of 3.623, with a significance level of *p*-value of 0.004, indicating high reliability in these assessments.

Academic integrity, with an ICC of 0.676, also indicated a moderate agreement. This outcome suggested that the assessors consistently identified the use of accurate and appropriate citations and the originality of the students' work. The measure comprised items such as the presence or absence of citations, the correct use of citation format and the detection of original work versus plagiarism, and these contributed to this level of agreement. The F-test value of 3.152 and a significance *p*-value of 0.008 affirmed this agreement.

In contrast, the 'critical thinking' skill had an ICC of 0.471 with a lower F-test value of 1.846 and a significance level of 0.095. These indicated that the reliability of the assessments for critical thinking, while suggestive of some consistency, was not statistically significant and may require a re-evaluation of the assessment criteria or methodology to enhance the reliability.

Despite some variances in individual categories, the overall inter-rater reliability for these combined transferable skills, with an ICC of 0.784, showed good inter-rater reliability across all categories when they were considered together. This result indicated that, while there may be subjectivity and individual interpretation in assessing each element, there was a general agreement in the overall evaluation of the students' essays. The F-test value of 8.651 and a highly significant *p*-value of less than 0.001 corroborated this, providing strong statistical evidence that the assessors' consensus was significant and reliable.

**Assessors' evaluation of the quality of student essays**

Table 3 used the Mann-Whitney U test to compare the transferable skills between the non-AI and AI cohorts. It allowed for a non-parametric comparison for sample sizes that were small, not normally distributed, and particularly useful for evaluating transferable skills that often

ranked (ordinal data) rather than measured on an interval scale. It also provided a reliable way to assess whether the observed skill differences were statistically significant.

**Table 3**

*Mann-Whitney U test comparing the transferable skills for the non-AI and AI cohorts*

| Transferable skills | Mann-Whitney *U* | Wilcoxon W | Z-score | Asymp. Sig. | Exact Sig. |
|---|---|---|---|---|---|
| Writing skills | 180.5 | 390.5 | -0.556 | 0.578 | 0.602 |
| Critical thinking | 93.0 | 303.0 | -3.025 | 0.002 | 0.003‡ |
| Information literacy | 98.0 | 308.0 | -2.920 | 0.003 | 0.005‡ |
| Academic integrity | 149.0 | 359.0 | -1.503 | 0.133 | 0.174 |
| Creative thinking | 137.0 | 347.0 | -1.917 | 0.055 | 0.091 |

‡ Significant difference between the paired observations, $p < 0.05$

The analysis of the independent assessors' assessment framework sheets showed that the 'writing skills' had a Mann-Whitney *U* value of 180.5 and a corresponding *p*-value of 0.578. This high *p*-value indicated no statistically significant difference in writing skills between the non-AI and AI cohorts. Similarly, 'academic integrity' and 'creative thinking' with *p*-values greater than 0.05 indicated no significant differences between the two cohorts for these skills.

Conversely, the Mann–Whitney *U* test showed significant differences between the non-AI and AI cohorts regarding critical thinking (Mann–Whitney $U = 93.0$, $p = 0.003$) and information literacy (Mann–Whitney $U = 98.0$, $p = 0.005$). This result suggested that the interventions or teaching methods applied to the AI cohort may have positively affected the development of students' critical thinking and information literacy skills compared with the non-AI cohort.

**Interpreting evidence of student learning through AI Interaction**

Integrating digital information and learning tools, such as Google Scholar and ChatGPT, into the learning process created a unique opportunity for educators to gather and interpret evidence of student learning. Here, we analysed the accessors' report and students' essays. The main themes were identified from the assessors' feedback and students submitted reports.

***Theme 1: Learning patterns and skill development***

Examination of the assessors' feedback showed that for the non-AI cohort, the main distinguishing feature between the higher-graded assignments and the lower-graded assignments was that higher-graded assignments generally showed a deeper understanding of technical knowledge and more awareness of the broader societal and environmental implications of their research topic. This outcome was similar to the feedback for the AI cohort, with the added observation that the higher graded assignments were notably better structured. This improvement may reflect the influence of a more comprehensive search strategy that was enabled through the use of the AI:

> *Some students, like in samples 2 and 4, touch on environmental or societal implications of their subjects, while others do not. This indicates a variation in the breadth of perspectives considered when researching and writing. The most notable different is the use of AI tools in Cohort 2 (AI cohort). This adds an extra dimension to their research and presentation, which is absent in Cohort 1 (non-AI cohort).* [Assessor #2]

Comparing the non-AI and AI cohorts within each grade level showed that the AI cohort used a more diverse range of references. This observation was similar to that in higher-graded assignments, where, again, the assessors commented on the diversity and depth of the content of the assignments:

> *Cohort 2 (AI cohort) HD students show innovation not only in their use of technology, but also in how they integrate this with traditional research methods, possibly leading to richer and more diverse content.* [Assessor #1]

By examining samples from the AI cohort in more detail, it can be seen that the student's approach to research began by querying ChatGPT to ensure they had a basic understanding of the main concepts associated with their chosen topic in Materials Engineering. Following this, they utilised online search engines to investigate the subject further. Their iterative approach, moving back and forth between ChatGPT and online searches, enabled them to cross-reference information and highlighted their developing skills in self-directed research. Such practices demonstrated that a student was not just passively consuming information but

actively engaging in learning, clearly shown in the insights of the assessors' excerpt and the students' work (Fig. 1).

> *The student started with a simple search to understand what rutile and ilmenite, the two major minerals of titanium, are, then refine their search strategy to find out different ways to refine titanium. This demonstrates a logical flow of idea and helped them to develop their coherent arguments.* [Assessor #1]

Some students drew comprehensive mind maps of how they used different sources of information and various search terms. In contrast, others made general reflective statements on how reliable they found the information from the AI tool and cited when they had to cross-check information (Fig. 1).

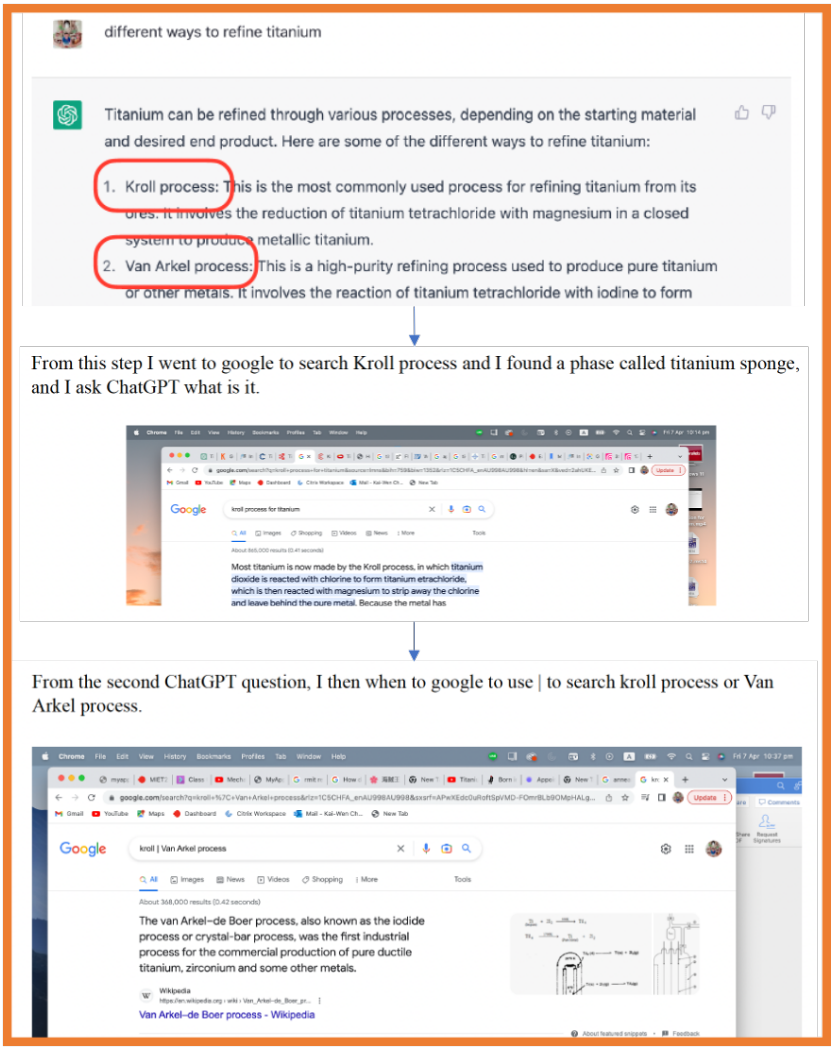

**Fig. 1**

*Students' methodical approach involved alternating between various resources for knowledge acquisition.*

***Theme 2: Area of improvement***

In the non-AI cohort, the lower-graded assignments generally lacked critical analysis and synthesis of information. The same observation was made in the AI cohort, especially those at the lower-grade level. Notably, a more diverse range of sources were used, and that, possibly, ChatGPT was mainly beneficial in the initial stages of the data search, as stated by one of the assessors:

> *Source diversity. Cohort 2 (AI cohort) had access to a more diverse range of sources due to the AI tool's ability to pull information from a vast database, whereas Cohort 1 relied more on traditional research methods.* [Assessor #2]

Among higher-graded assessments, the AI cohort showed more skill in analysing and synthesising the information. Examining this last observation in more detail, it was evident that the AI cohort exhibited good proficiency in analysis and synthesis, demonstrated by their good use of multiple sources to support the comprehension of their chosen topic. While students might excel in consolidating information from various resources with the help of AI, they need to foster their capacity for independent critical thinking. Students should be encouraged to dig deeper into topics, question underlying principles, and develop a more holistic understanding of the subject matter. An example of the assessor's excerpt:

> *The student provided search tactics to demonstrate their approach to understand the topic in depth. Despite, the student has demonstrated their ability to find information, perhaps still lacking the ability to synthesis the information for developing their own unique perspectives or hypotheses.* [Assessor #1]

***Staff reflection on assignment design and pedagogical strategies***

We investigated further into a reflective practice undertaken by the educator who teaches the subject—drawing from their extensive personal experience as a materials engineer. The educator's insights informed the design of a research assignment that aimed to cultivate the

capabilities of Aerospace, Mechanical, and Mechatronics Engineering students about their ability to source and appraise reliable material-related information efficiently. The educator's broader reflection on the pedagogical strategies necessary to cultivate critical skills in engineering disciplines, particularly focusing on the integration of new technologies like AI in the educational process, was provided here:

> *I considered what would happen if I made the students use the AI as part of their process for finding information. I played with the AI myself and began to see some of its limitations, in terms of accuracy of data, relevance, and currency of data. I also realized that using ChatGPT, in conjunction with other search tools, was how I would use the AI and that this should be a goal for students as well. If I made their information searching techniques part of the assessment, then this would make their data searching visible, but it might also highlight how well they integrate knowledge from different sources and order it into a logical framework. Using ChatGPT as one source among many also gives students experience in determining the reliability of the information which they find.*
> *To support them in this I developed learning materials which examined the pros and cons of the AI, and which also revised basic search strategies such as using wild cards and Boolean operators.*
> *I altered the marking rubric so that they were graded on their search strategy, the questions they asked ChatGPT and whether they iterated their search strategy between a standard search engine and the AI.*

This reflection led to an appreciation of enhancing pedagogical approaches to develop transferable skills that are imperative for engineering students. It also prompted a critical evaluation of the usage of AI tools, such as Chat GPT, in educational settings.

## Discussion

Integrating AI technologies is no longer an optional enhancement but a vital component that resonates with modern educational needs. Therefore, this study aimed to evaluate how generative AI, especially ChatGPT, can be used to augment transferable skills among students and explore the methodologies and strategies that can be adopted to incorporate ChatGPT effectively into existing higher education teaching frameworks.

The themes discussed in this study were derived from qualitative data sources, including assessors' comments and reflective statements provided by students as part of their

coursework. These themes emerged independently of the quantitative analysis and were identified after the statistical results were finalised. This sequential approach ensured that the qualitative insights complemented, rather than influenced, the statistical findings. For example, while the statistical analysis revealed significant differences in critical thinking and information literacy between the AI and non-AI cohorts, the qualitative data provided depth to these results by highlighting how students understood utilising AI tools to diversify sources and refine research strategies.

**Skills development**

From our findings, it was evident that the use of ChatGPT could significantly impact students' development of transferable skills, notably in the areas of critical thinking and information literacy, which reflected both their ability to analyse and synthesise information from multiple sources and to use evidence to support their arguments effectively. The significant ICC and F-test values for information literacy indicated that the assessors consistently recognised the students' competencies in these areas. This consistency suggested that the students were able to engage with complex material, discern relevant information, and integrate it effectively into their written work.

This finding was substantiated by applying the Mann–Whitney *U* test, which is a non-parametric statistical test, making it suitable for data that do not follow a normal distribution. In this study, the scores for transferable skills across the non-AI and AI cohorts were not normally distributed, as confirmed by preliminary tests for normality (e.g., the Shapiro-Wilk test). The test was used to compare differences in medians between two independent groups and evaluate the impact of AI exposure (AI cohort vs non-AI cohort) on the development of transferable skills. Each transferable skill was assessed on a scale reflecting performance levels based on grading rubrics (Appendix A). Given the relatively small sample size (20 essays in total, 10 from each cohort), the Mann-Whitney *U* test is appropriate because it is less sensitive to sample size variability than parametric tests, such as the independent *t*-test.

The results of the Mann-Whitney *U* test revealed considerable differences between the AI cohort and the non-AI cohort.

The AI cohort exhibited improvements in critical thinking and information literacy. AI tools such as ChatGPT facilitated access to diverse sources, enabling students to draw on a broader knowledge base. As noted by Assessor #2, students in the AI cohort benefited from AI's ability to summarise, refine and organise information, resulting in a "more diverse range of sources" compared to traditional research methods in the non-AI cohort. Several students in the AI cohort demonstrated logical and systematic research strategies, such as starting with broad searches and progressively refining their queries (e.g., searching for the minerals rutile and ilmenite before exploring titanium refinement methods). This structured approach, as highlighted by Assessor #1, likely contributed to their ability to develop coherent arguments and demonstrate deeper critical analysis. HD students in the AI cohort showed innovation in integrating AI tools with traditional research methods. As Assessor #1 noted, this combination added "richness and diversity" to their content, allowing them to explore novel perspectives, including environmental and societal implications of their topics. This feedback suggested that the interventions or teaching strategies implemented in the AI cohort successfully enhanced the students' ability to find, evaluate, and use information effectively and contribute to academic discourse. Students from the AI cohort had deliberately selected and implemented critical thinking strategies to develop their ideas. They demonstrated their ability to develop research ideas and apply progressive levels of information literacy through their submitted essays. These outcomes aligned with the broader higher education goal of fostering autonomous learners who can think critically and apply knowledge in various contexts (Burns et al., 2022).

The analysis of students' assessments further highlighted how AI tools enhanced the ability of students to ask follow-up questions and inquire into the 'why' and 'how' of the subject matter. Students pointed to a greater engagement with the learning materials and a deeper analytical and evaluative approach to learning. The findings of this study reinforced those of the limited

number of earlier studies that examined academic writing or students' essays, namely, that ChatGPT did support literacy development (Buruk, 2023; Li & Xu, 2023; Liu et al., 2023; Nyale et al., 2024).

Creative thinking showed borderline significance but did not reach statistical significance. Students may also have used AI as a source of ideas rather than a tool to inspire novel approaches. While helpful in generating ideas, AI tools rely on pre-existing data and may not effectively foster out-of-the-box thinking. This constraint could limit their ability to inspire truly novel or innovative solutions, key components of creative thinking. Similarly, while AI can assist students with grammar, coherence and structure, writing skills did not significantly improve in the AI cohort. Students may not effectively address deeper aspects of academic writing, such as developing unique arguments or synthesising diverse sources into a cohesive narrative. As Assessor #1 noted, some students demonstrated the ability to find information but "lacked the ability to synthesise it for developing their own unique perspectives or hypotheses.".

While our findings confirmed previous studies where improvements could be associated with the use of ChatGPT (Ellikkal & Rajamohan, 2024; Song & Song, 2023; Yilmaz & Karaoglan Yilmaz, 2023), they also indicate that such improvements do not necessarily improve learning outcomes without a more objective measure of improvement being developed and integrated into educator's pedagogical strategies. For instance, academic integrity did not show significant differences between the cohorts. As the study did not monitor how students in the AI cohort used ChatGPT, their usage could be inconsistent or opportunistic. While students were encouraged to document their search strategies, the absence of a controlled environment may have introduced variability in how academic integrity was practised or reported. In their reports, students identified the search engines and strategies they used, and this was valuable for the educators to know. However, such information may not accurately reflect actual skill development or changes in learning behaviour. Consequently, our study's findings do not necessarily reflect any direct causation in ChatGPT usage to enhance skills,

as the AI tools might have worked more as aids in information presentation rather than developing underlying cognitive abilities. The absence of detailed data tracking (such as the number of searches conducted, interaction frequency with the AI, or the depth of questions posed) meant that, while our findings are highly suggestive, we cannot conclude definitively that these learning behaviours changed as a result of the AI intervention alone. Future research should aim to develop and incorporate various metrics that provide more comprehensive analyses of how interactions with AI influence learning processes.

We acknowledged from the previous studies (Essien et al., 2024; Kaplan-Rakowski et al., 2023; Song & Song, 2023) that familiarity with AI tools could have influenced their ability to use ChatGPT effectively. While ChatGPT gained popularity quickly, its integration into educational contexts was still in its early stages at the time of the study, limiting the likelihood of substantial prior use among students. Students in the AI cohort were introduced to ChatGPT for this study, which included essential guidance on its usage. This structured introduction likely served as their primary exposure to the tool. It is important to acknowledge that students in the AI cohort may have had varying levels of exposure to other digital tools (e.g., Grammarly, citation managers, or search engines) and may have had an advantage in navigating ChatGPT's interface and leveraging its capabilities. It is also important to consider that regional and cultural contexts (reflecting differences in access to technology, educational priorities, and societal values) play a critical role in shaping how students perceive and utilise AI tools, as highlighted previously (Baidoo-Anu et al., 2024; Johnston et al., 2024; Kim et al., 2024). Students who come from regions with more advanced digital infrastructures may be more adept at using AI due to broader exposure to technology in their earlier education. Conversely, in areas where access to technology is limited, students might rely on traditional methods, influencing their engagement with AI tools. Students were encouraged to document their search strategies and interaction with ChatGPT to account for these external factors. This qualitative data suggested that most students engaged directly with ChatGPT for their assignment, and their approaches were shaped by the tool's guidance rather than by pre-

existing expertise or regional factors. However, further research is necessary to explore how broader cultural and technological contexts influence AI adoption and effective use in education.

**Pedagogical strategies to incorporate ChatGPT**

Integrating AI-generated information such as that provided by ChatGPT into educational contexts presents both a transformative opportunity and a significant challenge for educators. The powerful capabilities of AI models such as ChatGPT can, if harnessed correctly, amplify student learning outcomes, particularly in critical thinking. However, we must undertake this journey with a strong commitment to uphold ethical standards and prevent plagiarism or cheating, as previously mentioned (Alin et al., 2023; Cotton et al., 2023; Lodge, 2024). It is necessary to stress the significance of recognising intellectual property and equipping students with the essential skills and comprehension to differentiate their original thoughts from those produced by AI. As we move forward in this AI-influenced era of education, striking the right balance between leveraging AI's benefits and maintaining academic integrity will be paramount. Our study shed light on some essential practices, notably from the instructional approach in which students were guided to develop their search strategies, including listing key terms, Boolean operators, and in queries submitted to ChatGPT. By instructing students also to explain their ChatGPT search strategies clearly, educators could better assess students' understanding of research processes. Such practices encouraged critical engagement with AI outputs and prompted students to critique the AI-generated content they received rather than accept it. Assessing such search strategies could promote academic integrity by enabling educators to track and validate the students' information-gathering process. Having students disclose the queries they submitted to ChatGPT also encouraged students to engage consciously and ethically with AI. This disclosure allowed educators to guide students in formulating more effective questions and using AI as a tool for inquiry rather than merely for information retrieval.

The essays were marked by the educator responsible for delivering the subject, who used a rubric that assessed various aspects such as content, structure, written expression, and use of figures, diagrams and referencing. Generally, students were able to source reliable basic information. While students' logical thinking skills were evident in how they organised their essays, their methods for assessing the reliability of information from ChatGPT remained unclear. Recognising the growing importance of AI in research and learning, the educator for the subject revised the assessment rubric so that students were also graded on their search strategies, the questions they asked ChatGPT and whether they iterated their search strategy between another standard search engine and AI. By making information-searching techniques a part of the assessment, the revised rubric sheds light on students' data-searching abilities. It gauges how effectively they integrate knowledge from various sources into their essays.

From the study, we discovered that, while students demonstrated the ability to find relevant literature through meticulous search strategies, they appeared to lack the requisite skills to analyse and synthesise information and form their own perspectives and independent opinions. Hence, we suggested that the integration of ChatGPT must be coupled with additional instructional strategies to cultivate these higher-order skills. As suggested by Alin et al. (2023) and Nyale et al. (2024), educators need to reconsider assessment design to promote higher-order thinking, thereby rewarding students for synthesis and analyses of ideas and removing the attraction of cheating.

Our findings also indicated that with the help of ChatGPT, students' work demonstrated a more organised and logical structure. This observation could be due to the generative nature of the AI tool, which tends to generate keywords and technical terminologies with high correlations to the theme of the assessments. Students in the AI cohort also demonstrated their proficiency in raising follow-up questions when encountering unfamiliar terminology, and they could ask clarifying questions. These relevant terms helped and guided students in conducting their search progressively, step-by-step, in an organised manner.

**Promoting reflective teaching and learning practices: staff and student**

This study utilised an approach to gauge students' interactions with AI tools. Participants in the AI cohort were required to document their search strategies and the queries they submitted to ChatGPT. This process was designed to provide insights into how students utilised AI to enhance their search strategies and refine their use of online information resources. By analysing how students refined their searches and utilised ChatGPT responses, assessors gained a nuanced understanding of the development of information literacy and critical thinking skills among the AI cohort. This insight suggests that the observed improvements in skills can be partially attributed to the study intervention, rather than solely to participants' prior familiarity with AI tools.

Students can also be guided to interpret their AI interaction data so that self-monitoring and reflective learning are practised. By critically analysing their engagement with ChatGPT, students could identify strengths and areas for improvement, guiding their personal development and reinforcing transferable skills such as self-awareness and problem-solving.

**Implications**

In teaching Aerospace, Mechanical and Mechatronics Engineering, it was crucial to enable students to source and assess reliable information from various platforms. Our findings suggested that students can utilise AI, such as ChatGPT, as a 'research assistant' to guide them in their initial phase of information gathering. To do this effectively, students need to learn how to differentiate between creditable and non-credible sources of information and understand the limitations of generative AI in terms of accuracy, relevance and currency of data. Students can also leverage the natural language capabilities of generative AI to offer them immediate feedback on written expression and structure, enabling them to engage in a more iterative and active learning process.

The inclusion of ChatGPT in this study was not merely a technologically driven decision; its inclusion was substantiated through staff reflection on their professional experiences. Staff had recognised that the technological skills imparted through exposure to AI platforms, such

as ChatGPT, could be invaluable in the students' future careers, further emphasising the study's relevance and utility.

The data collected on students' search strategies and AI interactions will enable the course coordinator to evaluate and refine future assessment strategies and requirements for integrating emerging AI-driven tools. While the use of AI in educational settings has previously demonstrated potential to strengthen student learning and cultivate transferable skills (e.g. Essien et al. (2024); Kamalov et al. (2023); Nikolic et al. (2024); Nyale et al. (2024); Song and Song (2023)), this study found there are also challenges, notably in plagiarism. However, this study mostly evidenced plagiarism in poor paraphrasing rather than deception or collusion. Hence, strategies need to be considered for students using ChatGPT to guide them correctly in paraphrasing and training them in the ethical use of AI tools. Such guidance will help students enhance their learning experience by providing additional support in generating ideas and feedback on draft work. In the face of rapidly evolving AI technology, the study recommends ongoing assessments to ensure that AI integrations are aligned with educational values and ethical academic standards.

**Limitations and future research**

Inter-rater reliability tests were conducted to ensure consistent and unbiased in the evaluation processes by the two independent accessors. This approach ensured that the assessment was not merely a subjective opinion of one assessor but a reflection of a more standardised consensus. The inter-rater reliability tests also provided our research team with insights into gaps in the evaluation criteria, thereby enabling refinement and improvement. The poor inter-rater reliability in categories such as critical thinking and creative thinking skills suggested that the two independent assessors’ understanding of these assessment criteria in the rubric was inconsistent. While there was a basic agreement between assessors in certain areas, improvement was needed to achieve a higher consistency in evaluating critical and creative thinking skills.

The findings of this study were drawn from a relatively small sample of essays. This approach allowed for a focused comparison of transferable skills between the AI and non-AI cohorts, ensuring the representation of high-performing students. In educational assessment, these ICC and F-test values were informative because they provided insight into which areas of student performance were assessed consistently and which areas might require more explicit guidelines or training for assessors to achieve more reliable ratings. The findings also revealed the complexity and subjectivity inherent in assessing nuanced skills such as critical and creative thinking, which had lower ICC values. This limitation in essay, assessor numbers and the variability in individual assessment scores suggested that while the overall evaluation (when all transferable skills were analysed together) demonstrated good reliability (ICC value between 0.75 and 0.9), the outcomes should be interpreted cautiously. Increasing the sample size could provide greater statistical power and enhance the generalisability of the findings, particularly when analysing trends across a broader range of performance levels. Further research with a more extensive collection of essays could provide a more substantial basis for assessing these crucial academic skills.

Additionally, the involvement of multiple reviewers would improve the reliability of the assessments by mitigating potential biases or inconsistencies in grading criteria. Expanding the sample size and reviewer pool will be critical for future research to build on the current study's findings, enabling a more comprehensive understanding of how AI tools influence transferable skills in diverse educational contexts. These enhancements will ensure the robustness and credibility of conclusions drawn about the impacts of AI in higher education.

The study did not use an experimental design, which would be essential to measure the specific contribution of ChatGPT tools isolated from other educational interventions, previous student experiences, or related variables. Therefore, any conclusions on skills development should be treated with caution. However, the findings of the cohort comparison used in this study are highly suggestive of a meaningful contribution to students' AI use. Future research should consider using a longitudinal design with pre- and post-test assessments to explore

the effects of ChatGPT and similar AI tools. Such an approach would assist in evaluating the impact of these technologies on student learning and skills development. Future research should also focus on developing robust guidelines for the use of AI, including the ethical implications, and to ensure the development of guidelines to promote responsible and effective use. Finally, there is a need to investigate the long-term impact on students' learning outcomes and potential adjustments required in assessment and evaluation practices.

There are not many studies so far that examine the topic that is the subject of this paper. Further research could deepen our understanding of student experiences if it can consider social differences such as gender, background, ethnicity and race, disability and social class. Such differences may be implicated in students' experiences of AI technologies and are worthy of further study.

## Conclusion

The proliferation of generative AI technologies in educational settings has opened new opportunities to enhance learner engagement and learning outcomes. Academic institutes and educators alike are committed to advancing the integration of these intelligent systems in educational settings. This study investigated the impact of AI-driven tools, specifically ChatGPT, in higher education on facilitating the development of transferable skills such as writing, critical thinking, information literacy and creative thinking among undergraduate Engineering students. Our results strongly suggested that students who utilised ChatGPT had better critical thinking and information literacy skills as independent assessors assessed using a specific assessment framework. By investigating the impact of AI tools on learning, this study contributes to the ongoing dialogue on the potential and limitations of integrating AI technologies into future educational practice, which will require a thoughtful and well-structured strategy. Such integration must balance the benefits associated with the usage of AI and upholding academic integrity. Students will need to be guided and supported in their

use of AI. The study also revealed a need for complementary pedagogical methods to develop higher-order thinking skills among learners.

## Acknowledgements

The author(s) disclose that they have no actual or perceived conflicts of interest. The authors disclose that they have not received any funding for this manuscript beyond resourcing for academic time at their respective university. The authors sincerely thank all the students for agreeing to include their essays in the study.

## References

Alavi, M., Biros, E., & Cleary, M. (2022). A primer of inter-rater reliability in clinical measurement studies: Pros and pitfalls. *Journal of Clinical Nursing*, *31*(23-24), e39-e42. https://doi.org/10.1111/jocn.16514

Alin, P., Arendt, A., & Gurell, S. (2023). Addressing cheating in virtual proctored examinations: toward a framework of relevant mitigation strategies. *Assessment & Evaluation in Higher Education*, *48*(3), 262-275. https://doi.org/10.1080/02602938.2022.2075317

Anastasia, O., Saini, A., Zapata, G., Searsmith, D., Cope, B., Kalantzis, M., Castro, V., Kourkoulou, T., Jones, J., da Silva, R. A., Whiting, J., & Nikoleta Polyxeni, K. (2023). Generative AI: Implications and Applications for Education. *arXiv preprint*. https://doi.org/10.48550/arXiv.2305.07605

Baidoo-Anu, D., Asamoah, D., Amoako, I., & Mahama, I. (2024). Exploring student perspectives on generative artificial intelligence in higher education learning. *Discover Education*, *3*(1), 98. https://doi.org/10.1007/s44217-024-00173-z

Barrett, A., & Pack, A. (2023). Not quite eye to A.I.: student and teacher perspectives on the use of generative artificial intelligence in the writing process. *International Journal of Educational Technology in Higher Education*, *20*(1), 59. https://doi.org/10.1186/s41239-023-00427-0

Batista, J., Mesquita, A., & Carnaz, G. (2024). Generative AI and Higher Education: Trends, Challenges, and Future Directions from a Systematic Literature Review. *Information (Basel)*, *15*(11), 676. https://doi.org/10.3390/info15110676

Burns, A., de Bruyn, L., & Wilson, S. (2022). A rubric approach to assessing information literacy competency in tertiary curricula. *Journal of university teaching & learning practice*, *20*(1). https://doi.org/10.53761/1.20.01.10

Buruk, O. O. (2023). Academic Writing with GPT-3.5: Reflections on Practices, Efficacy and Transparency. https://doi.org/10.48550/arxiv.2304.11079

Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*, *20*(1), 38. https://doi.org/10.1186/s41239-023-00408-3

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, *20*(1), 43. https://doi.org/10.1186/s41239-023-00411-8

Cooper, G. (2023). Examining Science Education in ChatGPT: An Exploratory Study of Generative Artificial Intelligence. *Journal of Science Education and Technology*, *32*(3), 444-452. https://doi.org/10.1007/s10956-023-10039-y

Cope, B., & Kalantzis, M. (2023). Platformed Learning: Reshaping Education in the Era of Learning Management Systems. *Varieties of Platformisation: Critical Perspectives on EdTech in Higher Education. London: Palgrave Macmillan*.

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*, 1-12. https://doi.org/10.1080/14703297.2023.2190148

Dai, W., Lin, J., Jin, F., Li, T., Tsai, Y.-S., Gašević, D., & Chen, G. (2023). Can Large Language Models Provide Feedback to Students? A Case Study on ChatGPT. *2023 IEEE International Conference on Advanced Learning Technologies (ICALT)*, 323-325. https://doi.org/10.1109/ICALT58122.2023.00100

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

Darvishi, A., Khosravi, H., Sadiq, S., Gašević, D., & Siemens, G. (2024). Impact of AI assistance on student agency. *Computers & Education*, *210*, 104967. https://doi.org/10.1016/j.compedu.2023.104967

Ellikkal, A., & Rajamohan, S. (2024). AI-enabled personalized learning: empowering management students for improving engagement and academic performance. *Vilakshan - XIMB Journal of Management*, *ahead-of-print*(ahead-of-print). https://doi.org/10.1108/XJM-02-2024-0023

Essien, A., Bukoye, O. T., O'Dea, X., & Kremantzis, M. (2024). The influence of AI text generators on critical thinking skills in UK business schools. *Studies in Higher Education*, *49*(5), 865-882. https://doi.org/10.1080/03075079.2024.2316881

Farazouli, A., Cerratto-Pargman, T., Bolander-Laksov, K., & McGrath, C. (2023). Hello GPT! Goodbye home examination? An exploratory study of AI chatbots impact on university teachers' assessment practices. *Assessment & Evaluation in Higher Education*, 1-13. https://doi.org/10.1080/02602938.2023.2241676

Gilson, A., Safranek, C. W., Huang, T., Socrates, V., Chi, L., Taylor, R. A., & Chartash, D. (2023). How Does ChatGPT Perform on the United States Medical Licensing

Examination? The Implications of Large Language Models for Medical Education and Knowledge Assessment. *JMIR Med Educ*, *9*, e45312. https://doi.org/10.2196/45312

Harris, D., Coleman, K., & Cook, P. J. (2022). Radical rubrics: implementing the critical and creative thinking general capability through an ecological approach. *The Australian Educational Researcher*. https://doi.org/10.1007/s13384-022-00521-8

Huq, A., & Gilbert, D. (2017). All the world's a stage: transforming entrepreneurship education through design thinking. *Education & training (London)*, *59*(2), 155-170. https://doi.org/10.1108/ET-12-2015-0111

Johnston, H., Wells, R. F., Shanks, E. M., Boey, T., & Parsons, B. N. (2024). Student perspectives on the use of generative artificial intelligence technologies in higher education. *International Journal for Educational Integrity*, *20*(1), 2. https://doi.org/10.1007/s40979-024-00149-4

Kamalov, F., Santandreu Calonge, D., & Gurrib, I. (2023). New Era of Artificial Intelligence in Education: Towards a Sustainable Multifaceted Revolution. *Sustainability*, *15*(16), 12451. https://www.mdpi.com/2071-1050/15/16/12451

Kaplan-Rakowski, R., Grotewold, K., Hartwick, P., & Papin, K. (2023). Generative AI and Teachers' Perspectives on Its Implementation in Education. *Journal of Interactive Learning Research*, *34*, 313-338. https://www.learntechlib.org/primary/p/222363/

Khan, R. A., Jawaid, M., Khan, A. R., & Sajjad, M. (2023). ChatGPT-Reshaping medical education and clinical management. *Pakistan journal of medical sciences*, *39*(2), 605-607. https://doi.org/10.12669/pjms.39.2.7653

Khlaif, Z. N., Ayyoub, A., Hamamra, B., Bensalem, E., Mitwally, M. A. A., Ayyoub, A., Hattab, M. K., & 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. https://www.mdpi.com/2227-7102/14/10/1090

Kim, J., Yu, S., Detrick, R., & Li, N. (2024). Exploring students' perspectives on Generative AI-assisted academic writing. *Education and Information Technologies*. https://doi.org/10.1007/s10639-024-12878-7

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

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155-163. https://doi.org/10.1016/j.jcm.2016.02.012

Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., Madriaga, M., Aggabao, R., Diaz-Candido, G., Maningo, J., & Tseng, V. (2023). Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language

models. *PLOS Digital Health*, *2*(2), e0000198. https://doi.org/10.1371/journal.pdig.0000198

Kurtz, G., Amzalag, M., Shaked, N., Zaguri, Y., Kohen-Vacs, D., Gal, E., Zailer, G., & Barak-Medina, E. (2024). Strategies for Integrating Generative AI into Higher Education: Navigating Challenges and Leveraging Opportunities. *Education Sciences*, *14*(5), 503. https://www.mdpi.com/2227-7102/14/5/503

Lane, H. C., McCalla, G., Looi, C.-K., & Bull, S. (2016). Preface to the IJAIED 25th Anniversary Issue, Part 2. *International Journal of Artificial Intelligence in Education*, *26*(2), 539-543. https://doi.org/10.1007/s40593-016-0109-9

Lee, D., Arnold, M., Srivastava, A., Plastow, K., Strelan, P., Ploeckl, F., Lekkas, D., & Palmer, E. (2024). The impact of generative AI on higher education learning and teaching: A study of educators' perspectives. *Computers and Education: Artificial Intelligence*, *6*, 100221. https://doi.org/10.1016/j.caeai.2024.100221

Li, Z., & Xu, Y. (2023). Designing a realistic peer-like embodied conversational agent for supporting children storytelling. https://doi.org/10.48550/arxiv.2304.09399

Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education*, *21*(2), 100790. https://doi.org/10.1016/j.ijme.2023.100790

Liu, Y., Zhang, Z., Zhang, W., Yue, S., Zhao, X., Cheng, X., Zhang, Y., & Hu, H. (2023). ArguGPT: evaluating, understanding and identifying argumentative essays generated by GPT models. https://doi.org/10.48550/arxiv.2304.07666

Lodge, J. M. (2024). *The evolving risk to academic integrity posed by generative artificial intelligence: Options for immediate action*.

Luckin, R. (2018). *Machine learning and human intelligence : the future of education for the 21st century*. UCL Institute of Education Press.

Lytras, M. D., Alkhaldi, A., Malik, S., Serban, A. C., & Aldosemani, T. (2024). The Artificial Intelligence (AI) Landscape in Higher Education (HE): Current Developments, Opportunities, and Threats. In M. D. Lytras, A. Alkhaldi, S. Malik, A. C. Serban, & T. Aldosemani (Eds.), *The Evolution of Artificial Intelligence in Higher Education* (pp. 1-10). Emerald Publishing Limited. https://doi.org/10.1108/978-1-83549-486-820241001

Mamabolo, A., & Myres, K. (2019). A detailed guide on converting qualitative data into quantitative entrepreneurial skills survey instrument. *Electronic journal of business research methods*, *17*(3), 102-117. https://doi.org/10.34190/JBRM.17.3.001

Markauskaite, L., Marrone, R., Poquet, O., Knight, S., Martinez-Maldonado, R., Howard, S., Tondeur, J., De Laat, M., Buckingham Shum, S., Gašević, D., & Siemens, G. (2022). Rethinking the entwinement between artificial intelligence and human learning: What capabilities do learners need for a world with AI? *Computers and Education: Artificial Intelligence*, *3*, 100056. https://doi.org/10.1016/j.caeai.2022.100056

Markel, J. M., Opferman, S. G., Landay, J. A., & Piech, C. (2023). *GPTeach: Interactive TA Training with GPT-based Students* Proceedings of the Tenth ACM Conference on Learning @ Scale, Copenhagen, Denmark. https://doi.org/10.1145/3573051.3593393

Martín-Raugh, M., Kell, H., Ling, G., Fishtein, D., & Yang, Z. (2023). Noncognitive skills and critical thinking predict undergraduate academic performance. *Assessment & Evaluation in Higher Education*, *48*(3), 350-361. https://doi.org/10.1080/02602938.2022.2073964

McGraw, K., & Wong, S. (1996). Forming Inferences About Some Intraclass Correlation Coefficients. *Psychological Methods*, *1*, 30-46. https://doi.org/10.1037/1082-989X.1.1.30

Megahed, F. M., Chen, Y.-J., Ferris, J. A., Knoth, S., & Jones-Farmer, L. A. (2023). How generative AI models such as ChatGPT can be (mis)used in SPC practice, education, and research? An exploratory study. *Quality Engineering*, 1-29. https://doi.org/10.1080/08982112.2023.2206479

Nguyen, A., Kremantzis, M. D., Essien, A., Petrounias, I., & Hosseini, S. (2024). Enhancing Student Engagement Through Artificial Intelligence (AI): Understanding the Basics, Opportunities, and Challenges. *Journal of University Teaching and Learning Practice*, *21*(6), 1-13. https://doi.org/10.53761/caraaq92

Nikolic, S., Sandison, C., Haque, R., Daniel, S., Grundy, S., Belkina, M., Lyden, S., Hassan, G. M., & Neal, P. (2024). ChatGPT, Copilot, Gemini, SciSpace and Wolfram versus higher education assessments: an updated multi-institutional study of the academic integrity impacts of Generative Artificial Intelligence (GenAI) on assessment, teaching and learning in engineering. *Australasian Journal of Engineering Education*, 1-28. https://doi.org/10.1080/22054952.2024.2372154

Nyale, D., Karume, S., & Kipkebut, A. (2024). A comprehensive analysis of the role of artificial intelligence in aligning tertiary institutions academic programs to the emerging digital enterprise. *Education and Information Technologies*. https://doi.org/10.1007/s10639-024-12743-7

Pang, T. Y., Kootsookos, A., & Cheng, C.-T. (2024). Artificial Intelligence Use in Feedback: A Qualitative Analysis. *Journal of University Teaching and Learning Practice*, *21*(06).

Paulsen, M. B. (2013). A Model of Critical Thinking in Higher Education. In (Vol. 30, pp. 41-92). Springer International Publishing AG. https://doi.org/10.1007/978-3-319-12835-1_2

Peres, R., Schreier, M., Schweidel, D., & Sorescu, A. (2023). On ChatGPT and beyond: How generative artificial intelligence may affect research, teaching, and practice. *International Journal of Research in Marketing*, *40*(2), 269-275. https://doi.org/https://doi.org/10.1016/j.ijresmar.2023.03.001

Qadir, J. (2023). Engineering Education in the Era of ChatGPT: Promise and Pitfalls of Generative AI for Education. *2023 IEEE Global Engineering Education Conference*

*(EDUCON)*, *Kuwait, Kuwait.*, 1-9. https://doi.org/10.1109/EDUCON54358.2023.10125121

Rahman, M. M., & Watanobe, Y. (2023). ChatGPT for Education and Research: Opportunities, Threats, and Strategies. *Applied Sciences*, *13*(9), 5783. https://www.mdpi.com/2076-3417/13/9/5783

Roe, J., Perkins, M., & Ruelle, D. (2024). Is GenAI the Future of Feedback? Understanding Student and Staff Perspectives on AI in Assessment. *Intelligent Technologies in Education*, Advanced Online Publication. https://doi.org/10.70770/rzzz6y35

Rosenberg, S., Heimler, R., & Morote, E.-S. (2012). Basic employability skills: a triangular design approach. *Education & training (London)*, *54*(1), 7-20. https://doi.org/10.1108/00400911211198869

Shively, K., Stith, K. M., & Rubenstein, L. D. (2018). Measuring What Matters: Assessing Creativity, Critical Thinking, and the Design Process. *Gifted Child Today*, *41*(3), 149-158. https://doi.org/10.1177/1076217518768361

Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students [Original Research]. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1260843

Stone, B. W. (2024). Generative AI in Higher Education: Uncertain Students, Ambiguous Use Cases, and Mercenary Perspectives. *Teaching of Psychology*, 00986283241305398. https://doi.org/10.1177/00986283241305398

Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning and Teaching*, *6*(1).

Tack, A., & Piech, C. (2022). The AI Teacher Test: Measuring the Pedagogical Ability of Blender and GPT-3 in Educational Dialogues. Proceedings of the 15th International Conference on Educational Data Mining, Durham, United Kingdom.

Timmerman, B. E. C., Strickland, D. C., Johnson, R. L., & Payne, J. R. (2011). Development of a 'universal' rubric for assessing undergraduates' scientific reasoning skills using scientific writing. *Assessment & Evaluation in Higher Education*, *36*(5), 509-547. https://doi.org/10.1080/02602930903540991

UNESCO. (2023). *Education in the age of artificial intelligence,* . The UNESCO Courier. https://unesdoc.unesco.org/notice?id=p::usmarcdef_0000387029_eng

Vogt, W. P., Elaine, R. V., Dianne, C. G., & Lynne, M. H. (2014). *Selecting the Right Analyses for Your Data : Quantitative, Qualitative, and Mixed Methods* [Book]. The Guilford Press. https://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=771500&site=ehost-live&authtype=sso&custid=s5445732

Wang, T., Lund, B. D., Marengo, A., Pagano, A., Mannuru, N. R., Teel, Z. A., & Pange, J. (2023). Exploring the Potential Impact of Artificial Intelligence (AI) on International Students in Higher Education: Generative AI, Chatbots, Analytics, and International Student Success. *Applied Sciences*, *13*(11), 6716. https://www.mdpi.com/2076-3417/13/11/6716

Welsh, T. S., & Wright, M. S. (2010). 1 - What is information literacy? In T. S. Welsh & M. S. Wright (Eds.), *Information Literacy in the Digital Age* (pp. 1-11). Chandos Publishing. https://doi.org/https://doi.org/10.1016/B978-1-84334-515-2.50001-9

Yilmaz, R., & Karaoglan Yilmaz, F. G. (2023). The effect of generative artificial intelligence (AI)-based tool use on students' computational thinking skills, programming self-efficacy and motivation. *Computers and Education: Artificial Intelligence*, *4*, 100147. https://doi.org/https://doi.org/10.1016/j.caeai.2023.100147

Zhao, R., Yunus, M. M., & M. Rafiq, K. R. (2023). The Impact of the Use of ChatGPT in Enhancing Students' Engagement and Learning Outcomes in Higher Education: A Review. *International journal of academic research in business and social sciences*, *13*(12). https://doi.org/10.6007/IJARBSS/v13-i12/20258

Zohny, H., McMillan, J., & King, M. (2023). Ethics of generative AI. *Journal of medical ethics*, *49*(2), 79-80. https://doi.org/10.1136/jme-2023-108909

# Appendix A

*Assessment framework used by independent assessor to evaluate student's essay in relation to their skills development*

Student essays-ID Assessor:

| Items for evaluation | Criteria | Performance rating | | | |
|---|---|---|---|---|---|
| **Writing skills**: Identify potential differences in writing quality in terms of the structure and content of the essays to evaluate the students' ability to identify and articulate arguments, consider alternative viewpoints, and draw valid conclusions. | Grammar, word usage and organisation facilitate the reader's understanding of the essay (Timmerman et al., 2011). | **Not addressed**<br>• Grammar and spelling errors detract from the meaning of the paper.•<br>• Word usage is frequently confused or incorrect<br>• Subheadings are not used or poorly used.<br>• Information is presented in in a haphazard way | **Novice**<br>• Grammar and spelling mistakes do not hinder the meaning of the paper.<br>• General word usage is appropriate, although use of technical language may have occasional mistakes.<br>• Subheadings are used and aid the reader somewhat.<br>• There is some evidence of an organizational strategy though it may have gaps or repetitions | **Intermediate**<br>• Grammar and spelling have few mistakes.<br>• Word usage is accurate and aids the reader's understanding.<br>• Distinct sections of the paper are delineated by informative subheadings.<br>• A clear organizational strategy is present with a logical progression of ideas. | **Proficient**<br>• Correct grammar and spelling.<br>• Word usage facilitates reader's understanding.<br>• Informative subheadings significantly aid reader's understanding.<br>• A clear organizational strategy is present with a logical progression of ideas. There is evidence of an active planning for presenting information; this paper is easier to read the most. |
| **Writing skills** Comments: | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| **Critical thinking**: Evaluate the coherence and logical flow of arguments presented in the essays, looking for evidence of strong and well-supported claims, and logical progression of ideas. | Analyse and evaluate an issue (Paulsen, 2013) | **Lower-level thinking**<br>• Interpreting<br>• Identifying assumptions<br>• Asking questions for clarification | **Higher-level thinking skills**<br>• Analysing claims<br>• Synthesizing claims<br>• Predicting | **Complex thinking skills**<br>• Evaluating arguments<br>• Reasoning<br>• Inference making<br>• Problem solving | **Thinking about thinking**<br>• Metacognition<br>• Self-regulation |
| **Critical thinking** Comments: | | | | | |
| **Information literacy**: Assess the students' ability to analyse and synthesise information from multiple sources, and to use evidence to support their arguments. | Have the ability to effectively find, evaluate and use information (Burns et al., 2022) | **Foundational: Acquisition (Unistructural / Multistructural)** Information acquisition skills and, understanding of ethics and acknowledgement of source | **Foundational: Application (Multistructural)** Provide multiple lines of evidence in context | **Advanced: Acquisition and Application** (Relational) Acquire, integrate and evaluate multiple lines of evidence in context | **Innovative and Abstract: Application. (Extended Abstract)** Autonomous information acquisition; Initiates research ideas; progressive level of information literacy applied to new contexts and situations through effective communication |
| **Information literacy** Comments: | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| **Academic integrity**: Identify the students' ability to provide evidence of accurate and appropriate citation, proper use of sources, and originality of work. | Accurate and Appropriate Citation | **Need improvement**<br><br>Frequently fails to provide accurate and appropriate citation of sources. Numerous citation errors or omissions. Incomplete or inconsistent reference list. Struggles to integrate sources effectively. Evident plagiarism. | **Fair**<br><br>Inconsistently demonstrates accurate and appropriate citation of sources. Some citation errors or omissions. In-text citations or reference list may have inconsistencies. Attempts to integrate sources, but some inconsistencies or lack of clarity. Limited plagiarism. | **Good**<br><br>Generally demonstrates accurate and appropriate citation of sources. In-text citations and reference list mostly complete and consistent. Appropriately integrates sources to support arguments and ideas. Minor instances of plagiarism. | **Excellent**<br><br>Consistently demonstrates accurate and appropriate citation of sources. Integrates sources effectively to support arguments and ideas. No instances of plagiarism. |
| **Academic integrity** Comments: | | | | | |
| Creative thinking: Consider the overall originality and creativity of the essays, looking for evidence that the students have engaged critically with the topic and produced work that is their own. | Specific creativity strategy (Shively et al., 2018) | **Novice**<br>Students randomly selected and implemented a creative thinking strategy, and/or they were unable to leverage the strategy to improve their ideas. | **Developing**<br>Students selected and implemented a creative thinking strategy to develop their ideas. They explained how the strategy supported their creativity. | **Expert**<br>Students deliberately selected and implemented a creative thinking strategy to develop their ideas. They explained how the strategy supported their creativity. | |
| **Creative thinking** Comments: | | | | | |