



# Artificial intelligence-based large language models and integrity of exams and assignments in higher education: the case of tourism courses

## Inteligência artificial baseada em modelos de linguagem de grande escala e integridade de provas e trabalhos no ensino superior: o caso dos cursos de turismo

Abdullah Ülkü

School of Tourism and Hotel Management, Harran University, Department of Tourism Guidance, Türkiye, [abdullahulku@harran.edu.tr](mailto:abdullahulku@harran.edu.tr)

Received: 14.08.2023; Revisions required: 14.09.2023; Accepted: 28.10.2023

### Abstract

There is an increasing concern regarding the potential misuse of ChatGPT-4 in compromising the integrity of examinations and assignments. This study aims to examine the capabilities of ChatGPT-4 in critical thinking abilities, whether it poses a threat to examinations and assignments in higher education, and create a discussion agenda on this issue. ChatGPT-4 was asked to generate, answer, and criticize questions in tourism marketing, tourism management, tourism economics, tourist guidance, and gastronomy. The answers were evaluated according to universal critical thinking standards. The findings obtained from this study showed that ChatGPT-4 had commendable competence in several critical thinking standards and could produce human-like texts. However, there are certain domains that might be improved to comply more effectively with the expectations and norms of academia. Educators were recommended to use comprehensive approaches that combine technological and educational techniques to address the issue of cheating enabled by tools, such as ChatGPT-4, during assessment and exam processes.

**Keywords:** Large Language Models (LLM), Artificial Intelligence, ChatGPT, Tourism Education, Exams, Integrity.

### Resumo

Há uma crescente preocupação em relação ao possível uso indevido do ChatGPT-4 comprometendo a integridade de exames e trabalhos acadêmicos. Este estudo visa examinar as capacidades do ChatGPT-4 nas capacidades de pensamento crítico, avaliar se ele representa uma ameaça para exames e trabalhos no ensino superior e criar uma agenda de discussão sobre esse problema. Solicitou-se ao ChatGPT-4 que gerasse, respondesse e criticasse perguntas nas áreas de marketing turístico, gestão do turismo, economia do turismo, informação turística e gastronomia. As respostas foram avaliadas de acordo com padrões universais de pensamento crítico. Os resultados deste estudo mostraram que o ChatGPT-4 apresentou competência admirável em vários padrões de pensamento crítico e conseguiu produzir textos semelhantes aos humanos. No entanto, há certos domínios que podem ser aperfeiçoados para atender mais efetivamente às expectativas e normas acadêmicas. Recomenda-se que os educadores adotem abordagens abrangentes que combinem técnicas tecnológicas e educacionais para lidar com o problema da fraude, facilitada por ferramentas como o ChatGPT-4, durante processos de avaliação e exames.

**Palavras-chave:** Modelos de Linguagem de Grande Escala (LLM), Inteligência Artificial, ChatGPT, Educação em Turismo, Exames, Integridade.

### 1. Introduction

The advent of digital transformation in education resulted in a new era characterized by enhanced accessibility and new autonomy (Latifah, Budiyanto & Saputro, 2022; Nieuwoudt, 2020). Students from different backgrounds now have the opportunity to attain their educational objectives without being restricted by geographical or time restrictions (Gerashchenko & Kovalev, 2021). However, the change is not without its challenges (Adzima, 2021; Montenegro-Rueda, Luque-de La Rosa, Sánchez-Serrano & Fernández-Cerero, 2021). The issue of integrity has become increasingly prominent with the widespread adoption of resources for educational instruction and assessment in universities (Comas-Forgas, Lancaster, Calvo-Sastre & Sureda-Negreet, 2021; Klijn, Mdaghri Alaoui & Vorsatz, 2022).

The intersection of technology and education has long been a subject of significant scholarly inquiry (Hashim, 2018; Rahm, 2023). Learning Management Systems (LMS), such as Moodle, have significantly transformed the learning process by providing individuals with a collaborative platform that facilitates new approaches to learning (David et al., 2022; Maslov, Nikou & Hansen, 2021). However, with the increasing use of assessments, there has been a growing concern over the

reliability of these platforms (Ababneh, Ahmed & Dedousis, 2022; Coghlan, Miller & Paterson, 2020; Duham, Al-mamory & Mahdi, 2022; Nguyen et al., 2022).

The integration of sophisticated technology, especially artificial intelligence (AI), into educational settings has further complicated this discourse (Barrientos, Cuadros, Alba & Cruz, 2021; Benvenuti et al., 2023). In the field of examinations and assignments, the ChatGPT-4, a sophisticated language model developed by OpenAI, distinguishes itself using its capacity to generate a narrative that exhibits human-like composition. This phenomenon presents both potential opportunities and ethical dilemmas (Susnjak, 2022).

Large Language Models (LLMs) have the potential to serve as a supplemental resource for a variety of assessments and academic tasks (Bernabei, Colabianchi, Falegnami & Costantino, 2023). Additionally, they can offer deeper perspectives on the learning process and customise evaluations according to students' particular knowledge, abilities, and experiences (Tsai, Ong & Chenet, 2023). In general, LLMs have the capacity to offer efficient ways of evaluating learning, creating examinations, replying to questions, and assessing performance (Swiecki et al., 2022).



The increased capabilities of ChatGPT may introduce new components to this discourse. The recognized impact of LLMs on enhancing education is well known (Essel et al., 2022; Naidu & Sevnarayan, 2023; Pérez, Daradoumis & Puig, 2020). However, concerns are emerging regarding its potential for misuse, especially in the context of assessments (AlAfnan, Dishari, Jovic & Lomidze, 2023; Crawford, Cowling & Allen, 2023; Oravec, 2023; Yau & Chan, 2023). The development of advanced models like ChatGPT raises significant concerns regarding the preservation of academic and exam integrity (Cotton, Cotton & Shipwayet, 2023; Currie, 2023; Eke, 2023; Elbanna & Armstrong, 2023) due to their ability to produce content that bypasses conventional plagiarism detection methods (Misra & Chandwar, 2023).

Studies about cheating on exams and assignments (Chirumamilla & Sindre, 2019; Holden, Norris & Kuhlmeier, 2021; Noorbehbahani, Mohammadi & Aminazadeh, 2022; Von Grunigen et al., 2018) show that there are several ways to cheat, such as searching the Internet, joining social networks, pretending to be another student, and wireless devices like earphones and microphones. The possibility that ChatGPT-4 may produce human-like messages and exhibit critical and intellectual thinking skills raises the concern that this chatbot could be a new cheating method for exams and assignments (Susnjak, 2022). In this direction, the research question was formed as follows: "Do Large Language Models based on Artificial Intelligence represent a threat to the integrity of exams and assignments in Higher Education?"

The development of LLMs, such as ChatGPT-4, gives rise to notable concerns due to their remarkable ability to produce narratives that closely mimic human writing ability. The current literature has examined the potential misuse of AI in the context of evaluations. Nevertheless, a noticeable lack of understanding exists regarding the particular results of expanded LLMs in this specific domain, their capacities, and their potential to establish critical thinking standards.

This study aims to investigate the impact of ChatGPT-4 on exams in the field of tourism and the possibility of using it as a cheating tool in these exams. Also, this study aims to provide useful information for educators. Specifically, it tries to shed light on the many capabilities and possible disadvantages of ChatGPT-4 when used in exams. By making a contribution to the scope of the ongoing discourse, this study enhances the existing body of literature on this topic. This study examines the capabilities and potential challenges associated with Language Model Models (LLMs), with an emphasis on ChatGPT-4, within the academic domain. By engaging in this process, this study addresses the current knowledge gap and provides insights into these models' use and potential misapplication in the context of educational assessments. The findings of this study provide practical suggestions for educators and institutions. The recommendations aim to limit the misusing of LLMs. This study also advises on the ethical considerations of using LLMs within the educational environment.

## 2. Literature Review

Educational environments and challenges have been changing rapidly. These changes have fostered a shift toward active learning and creative thinking. Technology creates new learning settings and methodologies. Students' capacity to succeed in different environments reflects their rising educational goals. Changing teachers' perspectives and reorganising student-teacher interactions promote new educational methods (Petrolo et al., 2023). These methods are digitalization and virtualization. They are used across various educational activities to enhance overall learning outcomes and expand the scope of educational resources (Hwang, 2018).

The emergence of communication technologies has fostered the rise of distance learning (Yorkovsky & Levenberg, 2022). COVID-19 accelerated this shift, prompting educational institutions to rapidly transition from traditional to online instruction, necessitating adjustments in scheduling, environment, and curriculum (Agasisti & Soncin, 2021). This digital shift in education, popularized further by COVID-19, has made online education a prevalent global method, primarily facilitated through the Internet and supported by devices like smartphones and tablets (Asaqli, 2020; Xie, Siau & Nah, 2020). Online education offers flexibility, allowing students to access lessons and resources anytime, anywhere, saving time and costs associated with campus commutes (Mellieon, Jr & Robinson, 2021). However, challenges exist. The absence of direct student-teacher interaction can complicate understanding, and the self-paced nature might diminish motivation. Connectivity issues and subpar instructional materials can impede learning, especially in rural areas (Ilynykh, 2021).

Students may use some cheating methods during online exams and other assignments (Noorbehbahani, Mohammadi & Aminazadeh, 2022). Students may search the Internet, join social networks, pretend to be another student, hide notes, and use mobile phones, scientific calculators, MP3 players with calculators, and wireless headphones and microphones. Also, cheating could include downloading papers and claiming them as one's own, accessing test materials without authorization, and asking other students for answers. Teachers and students believe online examinations are easier to cheat. The biggest risks are students working together or using forbidden tools during the test (Holden et al., 2021).

Lanier (2006) reported that 41.1% of students admitted to cheating in online criminal justice and legal studies courses, compared to 21.3% in face-to-face courses. King & Case (2014) found that 15% of undergraduate business students admitted to cheating on exams, averaging 2.9 times each semester. Watson & Sottile (2010) noted that 32.7% of students across academic areas admitted to cheating online. Dendir & Maxwell (2020) highlighted a higher likelihood of cheating in online classes due to the absence of direct supervision and ease of accessing unapproved resources. However, not all online students cheat, and online proctoring can mitigate dishonesty



risks. A Wiley (2020) survey of 800 global university educators found that 93% believed online exams offer more opportunities for cheating.

A new risk has emerged with using AI in exams and assignments: Using chatbots. Chatbots are software applications that communicate with people through text and speech on a certain topic or area. Chatbots offer many educational and training purposes. Recent advances in this subject indicate that communication with technology, whether through natural language or speech, is more feasible due to technical improvements and people's growing familiarity with engaging digital entities (Smutny & Schreiberova, 2020). Chatbots may be used for exams and assessments (Kooli, 2023); students can ask chatbots any exam question and answer that question (Dias, Kamdi, Gharat & Chudhari, 2019). One of the most developed chatbots is ChatGPT. Susnjak (2022) mentioned that ChatGPT demonstrates the ability to manifest critical thinking abilities and produce writing that closely resembles reality with less input. Consequently, this poses a possible risk to the authenticity of examinations, especially in higher education environments where these assessments are more common.

The advent of generative AI models, such as ChatGPT is an important advance in the constantly changing field of higher education. Although these tools offer new and innovative methods for teaching and learning, they create concerns about academic integrity, especially in the context of assignments and exams (Perera & Lankathilaka, 2023). The availability of ChatGPT enables those with limited knowledge of technology to make use of its capabilities, potentially bypassing the challenges associated with authentic learning. These models have the capability to rapidly generate a large amount of information, which presents challenges for monitoring during assessments. Moreover, their capacity to produce apparently original content presents a challenge for traditional plagiarism detection methods. The excessive dependence on AI has the potential to hinder the development of critical thinking skills while also presenting the hidden danger of AI models promoting or generating deceptive academic material. Educators often encounter challenges in distinguishing real student effort from content produced by artificial intelligence, which presents a complex and intricate situation (Gimpel et al., 2023).

LLMs can be used in distance learning or online exams, but they can also pose a threat to the integrity of assignments. The integration of AI into the field of higher education has caused a new era in ways of teaching (Carvalho et al., 2022). One of the major advances in this field is ChatGPT. This adaptable language model has the potential to significantly improve the academic experience for students by offering a variety of uses. ChatGPT can serve as a tool for students to improve the effectiveness of their literature search techniques when completing assignments. Individuals can get selected collections of significant and up-to-date academic papers by entering specific questions, improving their understanding of their selected subject matter (Wagner, Lukyanenko & Paré, 2022). However,

students might use the advantages provided by ChatGPT in a negative manner. After completing the literature review, students can choose to use ChatGPT as a means to produce a written copy of their academic assignments. The accessibility of ChatGPT could promote excessive reliance among students on the model for solutions, thereby decreasing their preference towards self-sufficient critical thinking and problem-solving (Jungherr, 2023; Putra, Rangka, Aminah & Aditama, 2023).

In the case that a significant number of students choose ChatGPT as a means of requesting assistance, it is possible that they might get standardised or similar responses in answer to their assignment questions. The procedure of homogeneity is an obstacle for educators in terms of detecting individual understanding and can potentially decrease the variety of diverse viewpoints, which is a fundamental aspect of academic discussion. According to Hill (2023), to prevent this, educators can provide an environment that encourages students to critically evaluate and correct potential biases and similarities present in the outputs produced by chatbots.

There is concern over the potential effects of ChatGPT on the academic integrity of assignments. The discussion and worry regarding their performance on many exams and assignments has generated significant attention and concern (Hill, 2023). Some studies have been conducted on how ChatGPT performs in examinations in higher education institutions and other fields. According to a recent study in the medical domain (Gilson et al., 2022), ChatGPT demonstrated performance equal to that of a third-year medical student who successfully passed a medical question-answer examination. In another medical study conducted by Kung et al. (2023), ChatGPT showed a performance of over 60%, meeting the established standard. Lin et al. (2023) found that ChatGPT exhibited superior performance compared to both ophthalmology residents and practising ophthalmologists.

A study in the domain of law was undertaken by Choi, Hickman, Monahan & Schwarcz (2023). ChatGPT was evaluated in four distinct final examinations at the University of Minnesota. ChatGPT answered over 95 multiple-choice questions and 12 essay questions. ChatGPT's scores were sufficient to meet the passing requirements.

Geerling, Mateer, Wooten & Damodaran, 2023 evaluated ChatGPT's performance on a well-recognised economic knowledge assessment used in the United States. This assessment primarily focuses on assessing individuals' comprehension at the principle level. ChatGPT achieved an accuracy rate of 63.3% in responding to microeconomics questions, while it demonstrated an accuracy rate of 86.7% in addressing macroeconomics queries.

In a study about tourism courses, the usage of ChatGPT has emerged as a significant concern concerning the trustworthiness and dependability of examinations (Göktaş, 2023). The system exhibited weaknesses in accuracy and verification stages when responding to assignment questions



since it offered inaccurate answers and referenced non-existent sources. Nevertheless, the system exhibited acceptable performance in multiple-choice questions, correctly answering 16 out of 20 questions. This outcome serves as a confirmation of its competence in assessing knowledge regarding multiple-choice tourism-related topics.

Newton & Xiomeriti (2023) examined 41 research papers about the performance of ChatGPT. These studies together used a sample of 46204 multiple-choice questions (MCQs) derived from 97 distinct question sets. The researchers discovered that versions of ChatGPT using GPT-4 had a high success rate in multiple-choice question (MCQ) examinations. The collection of 41 studies encompasses various academic disciplines, including medical, dentistry, computer science, life support, business, law, economics, mathematics, parasitology, physics, chemistry, social work, engineering, anatomy, and accounting. Based on the available empirical evidence, it can be said that ChatGPT exhibits considerable potential as an assessment tool in several academic disciplines, assignments and exams. The significant challenge in academic evaluations is characterised by its consistent performance, adaptability with various question forms, and domain-specific proficiencies.

According to Eke (2023), generative AI systems like ChatGPT pose a real threat to academic and exam integrity because they can make students more likely to give the system their writing projects. This could lead to copying and other forms of cheating in education. ChatGPT can write logical and interesting text, but it does not fully understand the complexities of human language and conversation, which means that the system could sometimes come up with the wrong information, give damaging directions or biased information, or have limited knowledge because of how it was trained.

The utilization of ChatGPT may have negative impacts on academic and exam integrity due to its potential to enable deceptive practices and cases of plagiarism. ChatGPT possesses the ability to produce replies to questions that display a resemblance to human writing, hence presenting a difficulty in distinguishing between text created by the machine and that generated by a human. The use of ChatGPT by students to generate homework or test responses might lead to instances of academic dishonesty (Cotton, Cotton & Shipway, 2023).

This study considers the most advanced ChatGPT model, ChatGPT-4. GPT-4, which was made available to the public in March 2023, is capable of accepting visual queries, can manage 32,000 tokens at the same time, and has 100 trillion parameters (Barrot, 2023), whereas chatGPT-3.5 has 175 billion (Zaitso & Jin, 2023). This study seeks to assess the cognitive abilities of ChatGPT-4, a recently constructed AI agent, as well as its capacity to generate human-like language in tourism exams. There has been a growing apprehension over the potential misuse of ChatGPT-4, as it has become more possible to use its capabilities for the purpose of cheating on examinations, trespassing upon the principles of academic honesty.

**3. Methodology**

This study presents an explanation of the technique employed to assess the critical and higher-order thinking abilities of ChatGPT. First, ChatGPT-4 was asked to create critical thinking questions for undergraduate students in “tourism marketing, tourism management, tourism economics, tourist guidance, and gastronomy”. The questions should include a scenario. In the second stage, ChatGPT-4 was asked to answer these questions. Finally, ChatGPT-4 was asked to evaluate the answers critically. The examples of the prompts entered for the three phases are shown in Table 1.

**Table 1 - Prompts to create questions, answers and criticism**

<b>Question</b>	Please ask a critical and difficult question with a scenario for undergraduate students in tourism marketing.
<b>Answer</b>	Please answer this question with examples and supporting information.
<b>Criticism</b>	Can you please critically evaluate the following response to the above question, listing its strengths and weaknesses as well as suggestions for improvement?

ChatGPT was evaluated regarding its potential for formulating and answering questions and thinking critically about these answers in

tourism courses. The answers were examined according to the critical thinking standards of Paul (2005) (Table 2).

**Table 2 - Critical thinking standards**

<b>Clarity</b>	Could you elaborate? Could you give me an example? Could you illustrate what you mean?
<b>Accuracy</b>	How could we check on that? How could we find out if that is true? How could we verify or test that?
<b>Precision</b>	Could you be more specific? Could you give me more details? Could you be more exact?



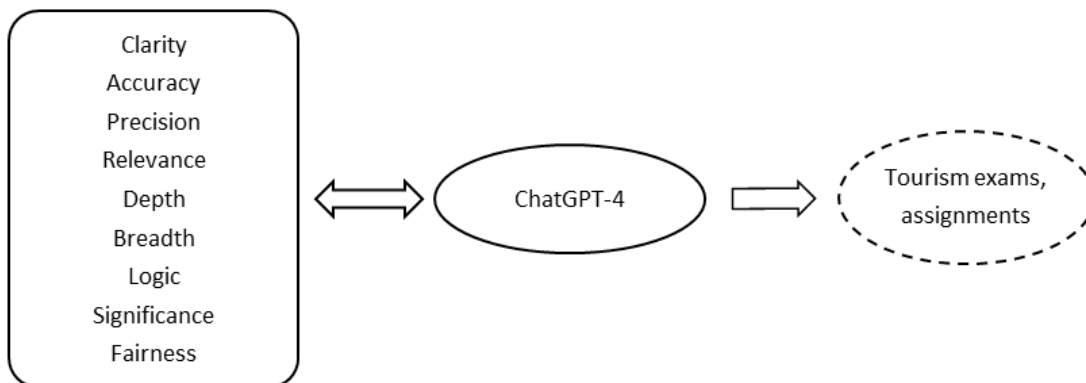
<b>Relevance</b>	How does that relate to the problem? How does that bear on the question? How does that help us with the issue?
<b>Depth</b>	What factors make this a difficult problem? What are some of the complexities of this question? What are some of the difficulties we need to deal with?
<b>Breadth</b>	Do we need to look at this from another perspective? Do we need to consider another point of view? Do we need to look at this in other ways?
<b>Logic</b>	Does all this make sense together? Does your first paragraph fit in with your last? Does what you say follow from the evidence?
<b>Significance</b>	Is this the most important problem to consider? Is this the central idea to focus on? Which of these facts are most important?
<b>Fairness</b>	Do I have any vested interest in this issue? Am I sympathetically representing the viewpoints of others?

Source: Adapted from Paul (2005).

Critical thinking is an idea that ties together what teachers and students need to know. It lets us help organizations change for the better. When critical thinking is known in the right way, it helps people internalize a deep knowledge of it. It is a skill that people need to be able to examine, evaluate, and understand

knowledge. It is a process of actively and skillfully conceptualizing, analyzing, synthesizing, and evaluating information to reach a conclusion or solve a problem (Paul, 2005). Figure 1 presents the possible critical thinking standards of ChatGPT-4.

Figure 1 - Possible critical thinking standards of ChatGPT-4 concerning tourism exams/assignments



According to Paul (2005), the standards of critical thinking and intellectuality are below:

**Clarity:** It means understanding without ambiguity or confusion. It means a statement is unambiguous. Because ambiguity makes it hard to judge a statement’s accuracy or relevance, clarity is a “gateway” criterion. Elaboration, illustration, and exemplification questions can improve comprehension (Paul & Elder, 2013).

**Accuracy:** The lack of faults, blunders, or distortions and the state of being truthful and accurate are intellectual standards of accuracy. It signifies truthfulness. Thinking is always accurate, so it is safe to assume a statement’s accuracy hasn’t been fully assessed until its authenticity is validated. One might ask questions about facts and sources to strengthen their reasoning.

**Precision:** The ability to be precise is an intelligence criterion. It means a comment is clear and precise. Unless one specifies a

statement, one may not fully understand it. One can ask particular questions to strengthen reasoning.

**Relevance:** It is an intellectual criterion that requires a connection to the topic. It must be significant and logically related to the issue. It means a comment is relevant to the situation. Thoughts can always stray from the work, question, problem, or issue at hand; thus, it is advisable to assume that individuals have not properly examined their thinking until they have considered all relevant topics, thoughts, and facts. Ask questions to make sure your thoughts are relevant.

**Depth:** It is considering all aspects of a situation, concept, or subject. It means considering all factors. A statement or idea examines a problem’s beliefs, implications, and impacts. Thinking can uncover what is underlying or work on the surface. Asking questions about a problem’s assumptions, meanings, and consequences helps deepen thinking.





**Breadth:** Academic breadth requires openness and multi-perspective thinking. It means a statement or thought process recognizes a problem's complexity and considers other viewpoints. Open-mindedness demands multi-perspective reasoning. Assessing how much thought it takes to understand a stream of thought is essential.

**Logic:** Logical things do not conflict with each other. It is founded on rationality. A remark or concept makes sense and does not conflict with itself. Finding out how rational thought might be beneficial. Asking questions to examine the consistency, coherence, and veracity of reasoning and proof will help think rationally.

**Significance:** It is the intellectual norm of being significant, impactful, or meaningful. It signifies that a remark or notion is significant, meaningful, and has significant implications. It is good to assess a thought's relevance to the situation. Questions about the most essential facts, ideas, and thoughts about a topic and what they mean or lead to are significant.

**Fairness:** Intellectual fairness excludes bias, dishonesty, partiality, self-interest, lying, and injustice. It means a statement or thought is fair, unbiased, and evaluates all important points of view without considering the speaker's feelings or interests. Asking questions encouraging people to explore alternative viewpoints and treat everyone equally, regardless of their feelings or goals, can help them think more fairly.

#### 4. Results

This section presents the evaluation of the question generation and critical thinking skills of ChatGPT. The answers to each question are examined separately.

##### 4.1 ChatGPT's capability for question generation

###### Tourism marketing

The question developed regarding Tropica's tourism marketing problem addresses the key elements of sustainable tourism, financial objectives, and the intricate "Tourist Life Cycle" framework, thereby demonstrating ChatGPT's understanding of the complex nature of tourism matters. Nevertheless, to further improve the depth of analysis, one might think about including empirical data-driven views or placing importance on the cultural dimensions of Tropica. The question presented by ChatGPT displays an important variety in its formulation. However, increasing its depth and relevance to academic or professional conversations could be achieved by adding more specific and data-oriented questions.

###### Tourism management

The question presented by ChatGPT on Marinelle's economic transition effectively underscores the inherent conflict between traditional means of subsistence and the emergence of new economic methods. The presented scenario exhibits a clear and cohesive depiction of the city's change and the subsequent difficulties that arise. While the question effectively

captures the socio-cultural details, it would benefit from a deeper exploration of specific management strategies to enhance precision. The importance of the topic is clear; yet, augmenting its complexity, including additional diverse aspects, such as environmental implications or worldwide trends in tourism, may prove beneficial. In addition, enhancing the scope of the analysis could be achieved by adding references to other global settings or considering the involvement of foreign actors, giving an extensive point of view. The movement from a comprehensive scenario to the search for answers demonstrates coherent logic. However, a more comprehensive understanding might be achieved by highlighting the intricate interplay of multiple components.

###### Tourism Economics

The question asked by ChatGPT regarding Econolia combines the fields of tourism and economic difficulties, demonstrating the complex relationships among several sectors. This question demonstrates a level of detail by effectively addressing both severe and underlying crises. The quality of the text is apparent in its careful presentation of different economic concepts, such as the multiplier effect. It may be beneficial to incorporate contemporary global trends or specific economic models related to tourism to enhance the relevance of future questions. Furthermore, although the given scenario addresses important economic considerations, it would be beneficial to delve into more complex implications, such as the socio-cultural consequences of economic transformations. In brief, the question demonstrates an extensive understanding of the field of tourism economics, while there is potential for further enhancement to make it more consistent with the complexities observed in real-world scenarios.

###### Tourist Guidance

The question produced regarding "Sylvantours" addresses the interaction of historical tourism with developing technology, namely augmented reality (AR) and virtual reality (VR). The text shows depth by emphasising a combination of cultural and technological fields and argues for a harmonic integration of conserving authenticity and embracing innovation. However, further developing the topic would be beneficial by describing the components of Augmented Reality (AR) and Virtual Reality (VR) that are most important to visitor help. Additionally, it would be useful to investigate the potential response from tourists visiting Techlandia regarding their use of AR and VR technologies. Although the current structure of the case is logical, it might be enhanced by including deeper links between the applications of augmented reality (AR) and virtual reality (VR) technologies and specific Ancienta landmarks, increasing the overall relevance of the discussion. The topic examines the key issues surrounding modern difficulties in tourist guidance. However, it may benefit from a deeper examination of the complexities associated with using technology in a heritage context.



## Gastronomy

The question presented addresses the contrast between Gastroville's existing culinary traditions and the emerging impact of molecular gastronomy. The question underlines the change in gastronomic preferences by drawing attention to the issues experienced by traditional restaurants due to the emergence of locations such as "MolecuEat." However, the approach's effectiveness could be enhanced by placing a higher priority on the socio-economic aspects of local enterprises or by delving further into the psychological attraction of contemporary vs traditional gastronomic encounters for customers. The current structure and clarity of the text capture the essence of the gastronomy debate. However, to further enhance clarity, it may be beneficial to provide a more explicit differentiation between the preservation of culinary traditions and the commercial sustainability of restaurants. In conclusion, the question sheds light on the dynamic culinary environment, although it would be beneficial to undertake a more complex examination of gastronomic identity and the associated economic consequences.

ChatGPT has shown an important skill level in formulating questions about many tourism aspects. It conveys key concepts ranging from sustainability to the integration of technology with heritage. The approach to tourist marketing employed by the organization is firmly based on fundamental ideas, such as sustainability and the concept of the "Tourist Life Cycle." In the field of tourism management, it regulates the complex harmony between traditional practices and rising economic opportunities. The model demonstrates an effective grasp of complex economic processes when applied to tourism economics. However, it might potentially benefit from a more nuanced understanding of global trends. The endeavour's exploration into providing visitor information demonstrates an appealing combination of technical progress and cultural authenticity. Within the domain of gastronomy, it illustrates the continuous development of culinary customs in the midst of continuous gastronomic changes. In conclusion, although ChatGPT demonstrates a basic understanding of tourism, there is room for improvement in enhancing its depth and nuance. This improvement would contribute to a better alignment between its outputs and the complexities observed in the real world.

## 4.2. Evaluation of Critical Thinking Skills of ChatGPT

### Tourism marketing

*Clarity* was noticed through the use of systematic explanations and the clear distinction between the problems at present and possible solutions. GPT-4 maintained high *precision* despite the limited practical use of the hypothetical situations. The solution showed precision by efficiently recognising the problems associated with uncontrolled tourism growth and providing sustainable strategies to address them. However, an additional review of the incorporation of local customs might have improved the analysis. The discussion centred on the issues faced by Tropica, with occasional angles exploring bigger global

trends that lacked a clear *relevance*. Regarding *depth*, the response properly recognized the complex issues involved in balancing the economic benefits of tourism with its environmental and social consequences. However, a more comprehensive examination of the fundamental factors at play would have enhanced the overall discourse. *Breadth* was maintained by including a wide range of stakeholders and their perspectives. The narrative exhibited *rationality* as it systematically explored the advantages, disadvantages, and finally proposed balanced resolutions related to the field of tourism. The *significance* of implementing sustainable practices in tourist locations, such as Tropica, was emphasized, indicating that a more thorough investigation of the specific choices made by Tropica may have enriched the discourse. Finally, the discourse demonstrated a commitment to *fairness* by trying to achieve an equitable balance between the economic benefits derived from tourism and the potential negative consequences associated with it.

### Tourism Management

A clear sense of clarity characterised the difference between Marinelle's historical significance and contemporary tourist demands. The importance of empirical information to support arguments was emphasised by *accuracy*, while *precision* argued for using specific examples to strengthen the suggested methods. The concept of *relevance* is centred on the equilibrium between traditional principles and current goals, although it would be beneficial to link it further with Marinelle's specific characteristics. The review of cultural traditions was conducted in a detailed way, considering the shifting preferences of tourists. This analysis suggests the possibility of conducting a deeper investigation. The *breadth* of the analysis commendably covers a wide range of viewpoints, including those of local residents as well as international tourists. The issue of integrating tradition with modernity was emphasised by *logic*, which suggests the need for a deeper examination of cause-and-effect relationships. The *significance* of Marinelle's transformation was highlighted, emphasizing the broader implications. However, it is suggested that a more detailed exploration of specific situations might provide more understanding. *Fairness* finally highlighted a balanced consideration of both traditional norms and progressive perspectives, arguing for a more impartial and well-rounded discussion.

### Tourism Economics

The emphasis on *clarity* was achieved by simplifying complex subjects, hence ensuring the accessibility of the text. The *accuracy* improvement was achieved by including hypothetical scenarios in real economic circumstances, further supported by empirical evidence derived from economic research. The issue of *precision* was examined in relation to certain economic challenges, although there was a suggestion for a deeper look into areas of impact and the diversification of businesses. The *significance* of tourism's influence on the economy of Econolia



was underlined. The evaluation examined the immediate and potential future implications of tourist dependence in relation to its *depth*. The *breadth* of this study underlined an extensive and varied viewpoint while also suggesting the need for additional investigation into visitor attitudes and patterns of purchase. The current dual-economic situation of Econolia and its associated consequences were brought to attention through *logical* examination. The emphasized *significance* of Econolia's problems lies in their wider implications for economies heavily relying on tourism. Finally, the concept of *fairness* supports adopting an equal position while examining tourism economics' potential drawbacks and benefits. It recommends incorporating various viewpoints to ensure a comprehensive and well-rounded evaluation.

### **Tourist Guidance**

Achieving *clarity* was made easier through the separation of traditional attractions and emerging technologies. The importance of *accuracy* was made clear by the practical growth of tourist information in real-world settings. This answer aims to explain the advantages and disadvantages associated with integrating technology in the context of historical tourism. The main goal of the *relevance* aspect was to ensure that the augmented reality (AR) and virtual reality (VR) apps maintained the cultural and historical integrity of Ancienta. The answer performed an analysis of the complex influence of technology on cultural tourism while also suggesting a need for more research into the fundamental drivers behind the use of technology. The *breadth* of the answer showed a diverse range of perspectives while also highlighting possible paths for further research, such as the examination of international case studies. The issue of properly integrating emerging technology with a deep historical backdrop is highlighted by logic. The *significance* of technology was emphasised, indicating the need for a deeper look into its socio-cultural, economic, and perceptual consequences. Finally, the principle of *fairness* promoted a comprehensive assessment of the advantages and disadvantages of implementing augmented reality (AR) and virtual reality (VR) technologies within a historically significant setting.

### **Gastronomy**

Achieving *clarity* was realized by deliberately combining historical culinary traditions with contemporary advancements, ensuring accessibility to individuals who might not have a particular interest in gastronomy. The primary concern of accuracy is in the real difficulties that occur when attempting to reconcile authenticity and innovation. The answer emphasised the use of contemporary methodologies, specifically molecular gastronomy, in the context of traditional culinary arrangements, although it could have provided a deeper examination of its implications for culinary education. The importance of *relevance* was underscored through the emphasis on modern techniques employed in preserving cooking traditions. The answer offered an examination of the consequences associated with the integration of traditional

practices and innovative approaches. However, it was recommended that a deeper examination of the fundamental ideologies involved be conducted. *Breadth* acknowledged the response's ability to present a well-rounded viewpoint while suggesting the advantages of integrating global contrasts. The logical analysis presented a cohesive representation of the confluence between the historical gastronomy of Gastroville and modern culinary skills. The *significance* of this culinary junction was discussed, highlighting its global relevance and suggesting the need for additional comprehensive research. Lastly, the principle of fairness underscores the significance of adopting a neutral approach, arguing for the incorporation of a wide range of opinions to achieve a comprehensive and balanced perspective.

The critical thinking skills of ChatGPT show themselves when evaluated in different facets of tourism, as it consistently suggests a strong dedication to clarity, precision, and depth. Within the domain of tourism marketing, ChatGPT shows an effective ability to evaluate and break down the complexity associated with sustainable practices. However, a deeper investigation of the nuances related to local traditions could enhance the depth and quality of its analysis. When examining Marinelle's tourism management, it skillfully negotiates the delicate balance between maintaining its historical heritage and achieving present-day aims. However, a heightened emphasis on complexity unusual to the local context could enhance the depth and quality of its insights. Regarding the tourism economics of Econolia, ChatGPT effectively simplifies complex economic challenges. However, a deeper exploration of the various aspects of economic diversification could enhance the depth of analysis. When examining tourist guidance, ChatGPT effectively identifies the conflict between traditional attractions and the increasing influence of modern technology, emphasizing the importance of thoroughly examining the root causes that lead to the use of technology. Within gastronomy, a harmonious combination of traditional culinary customs and modern innovative techniques exists. However, a deeper examination of the underlying ideas that support various culinary art forms could provide valuable insights. In summary, although ChatGPT provides a broad overview of various perspectives, a more focused examination of local-specific facts and fundamental philosophies could enhance the depth and precision of its conversation.

### **5. Discussion**

There has been a focus (Susnjak, 2022) on the value of critical thinking in both academic and practical contexts. The ability to objectively examine information holds significant relevance, especially in the contemporary world, where the Internet facilitates the rapid dissemination of incorrect information. Chatbots have the capacity to bring about a crucial change in various domains, such as the scientific realm, by expediting the distribution of knowledge. ChatGPT can respond to questions, provide informative cases, and join discussions with individuals seeking knowledge.

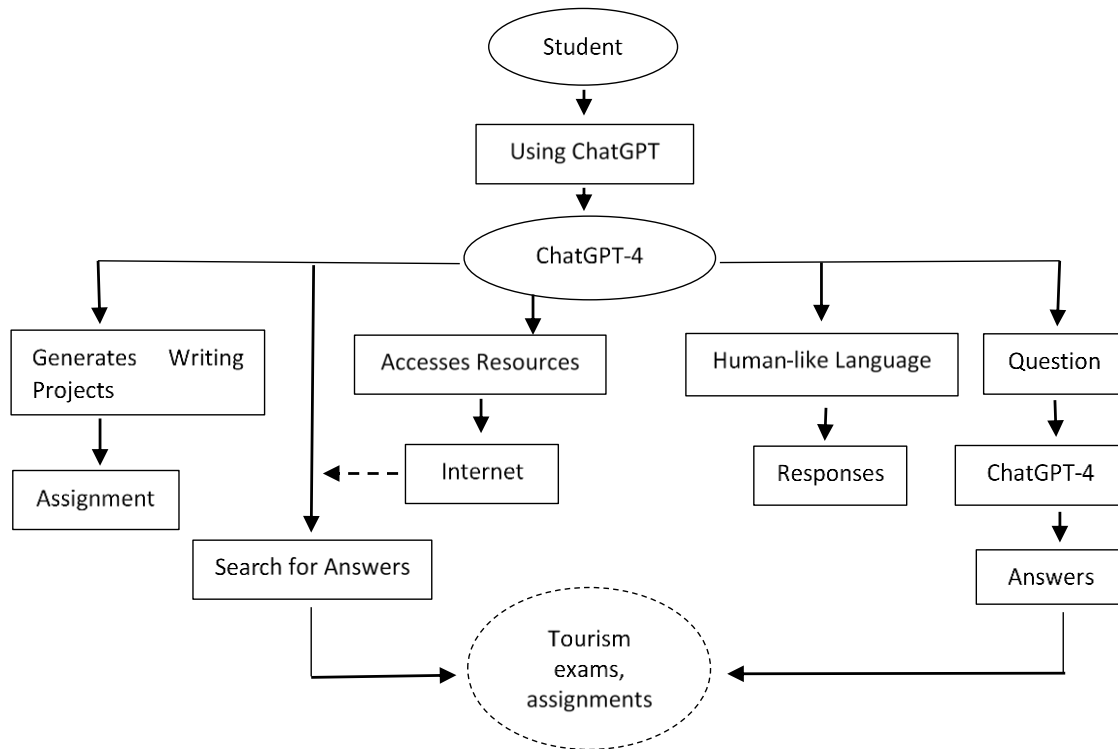




ChatGPT can be utilized for evaluation objectives, such as evaluating the difficulty level in generating assessments. According to Ventayen (2023), the utilization of AI in academic settings raises ethical considerations concerning academic and exam dishonesty issues. ChatGPT can write text that is nearly indistinguishable from content written by humans. This can make it challenging for tools like Turnitin to find plagiarism. Concerns have been raised that students might use ChatGPT to

cheat on tests or assignments (Skavronskaya, Hadinejad & Cotterell, 2023). Figure 2 shows the potential features that ChatGPT-4 can provide in tourism exams or assignments. The use of ChatGPT as a means of academic dishonesty may occur when students overly depend on it to provide solutions rather than employing it as a supplementary resource that enhances their learning experience.

**Figure 2 - Using ChatGPT-4 for tourism exams and assignments**



**Source:** Created by the author.

ChatGPT has several significant strengths. It uses an advanced natural language model to produce logical and coherent answers. Additionally, it has the capacity to enhance its own performance over time. ChatGPT exhibits some drawbacks, a deficiency in deep understanding, and challenges in assessing the quality of its generated responses. The difficulties faced by the field of education include a deficiency in threatening the principles of academic integrity and enabling the widespread occurrence of plagiarism. The findings presented in a study conducted by Farrokhnia, Banihashem, Noroozi & Wals (2023) support this study's results.

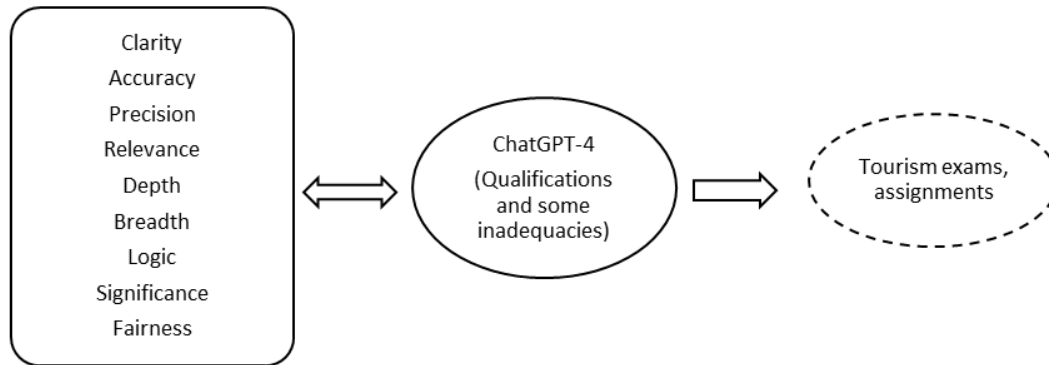
According to the results of this study, ChatGPT-4 demonstrates notable capabilities in several intellectual criteria; nonetheless, there are certain aspects that could benefit from further improvement to better accord with academic norms and expectations (Figure 3). This view is consistent with the worries expressed by scholars in the wider academic field about the potential risks associated with too much reliance on AI for important tasks in the

absence of human supervision (Cotton et al., 2023; Currie, 2023; Eke, 2023).

The ethical use of ChatGPT in examinations depends on how it is integrated into the assessment process and how students use it (Elbanna & Armstrong, 2023). Educators need to ensure that students have a comprehensive understanding of the limitations associated with ChatGPT and that its use is closely supervised to reduce cases of academic dishonesty. ChatGPT-4, an AI language model, can potentially facilitate academic dishonesty by enabling cheating during examinations or assignments. Several studies (Ali, Barhom, Marino & Duggalet, 2023; Atlas, 2023; Currie, 2023; Geerling et al., 2023) support this view. The utilization of ChatGPT-4 can serve both positive and negative objectives, and students who employ it without permission or in unethical ways are breaking the principles of academic honesty. OpenAI, the organization responsible for the development of ChatGPT-4, is currently engaged in advancing technology aimed at facilitating the identification of text produced by ChatGPT-4.



Figure 3 - ChatGPT-4 as a tool for tourism exams and assignments



Source: Created by the author.

Consequently, educators are advised to be proactive and ethical when utilizing these tools. In summary, although ChatGPT-4 possesses the potential to facilitate academic dishonesty during examinations and assignments, educational professionals and institutions should implement proactive methods to counteract this issue and uphold the principles of academic integrity.

## 6. Conclusion

The rapid growth of AI, particularly in sophisticated language models, such as ChatGPT-4, has caused intense discussion regarding its implications for several fields, including tourism education. The aim of this study within the field of tourism education was to assess the capacity of ChatGPT-4 to generate and assess queries that require critical thinking across several subdisciplines of tourism. The present study uses the critical thinking standards suggested by Paul (2005) to conduct a thorough assessment of the model's usefulness.

When evaluating the performance of ChatGPT-4 in relation to the critical thinking standards established by Paul (2005), many significant findings come to light. ChatGPT-4 effectively conveys concepts precisely, demonstrating a limited degree of complexity. Although the model's accuracy is often reliable, it occasionally relies on hypothetical circumstances that may lack practical significance. While ChatGPT-4 provides accurate replies, there is room for improvement in terms of enriching the depth of its content. The text primarily emphasises the importance of relevance, with intermittent deviations. Although many responses demonstrate depth, certain situations necessitate a more nuanced examination.

The model demonstrates breadth by incorporating many opinions, although there is potential for greater diversity. The logical coherence exhibited by ChatGPT-4 is a notable feature, as it guarantees a systematic and organized development of ideas. The examination of the key concepts or the significance of the subject matter has been partially accomplished, while a more comprehensive investigation into wider implications may prove beneficial. Finally, with regard to fairness, the model tries to balance its replies; however, a more comprehensive depiction of all viewpoints would improve its dedication to this criterion.

ChatGPT-4 model shows competence in various aspects of critical thinking, including clarity, accuracy, precision, relevance, depth, breadth, logic, and fairness. However, there are still areas that could benefit from further improvement. The expertise shown by artificial intelligence-based LLM, such as ChatGPT-4, may be posed as a reasonable concern regarding the preservation of the authenticity of examinations and academic tasks within the realm of higher education. Students may misuse their capacity to provide clear, logical, and pertinent feedback that showcases critical thinking to improve the quality of their academic assignments artificially. Students' potential use of AI for completing assignments or providing exam responses can hinder real-world education and the development of critical thinking skills and compromise the fundamental objectives of educational assessments.

The results of this study are significant for education. According to the results, it is highly likely that the use of ChatGPT in exams and assignments for marketing, management, economics, tour guiding, and gastronomy, which are important areas of tourism education, will have significant effects. Consequently, the advent of AI technologies, such as ChatGPT-4 can offer advantages and potential risks for tourism exams and assignments. According to the author of this study, it becomes apparent that although AI models have achieved notable advancements in natural language processing, they are not devoid of errors.

## Theoretical Implications

This study intends to fill the gap in the existing literature by examining the ethical implications and issues associated with the integration of AI in the field of education, specifically in relation to the comparison between conventional cheating tactics and AI-related risks.

The study highlights an important shift in our understanding of academic honesty within today's environment. The growing sophistication of AI and other digital technologies, demonstrated by ChatGPT-4, presents complex problems that pose a challenge to traditional understandings of cheating and academic integrity. The process of evolution in academia calls for researchers to seriously examine the fundamental principles that define academic integrity in the modern era. Moreover,



the study investigates the thinking skills of artificial intelligence, employing ChatGPT-4's effectiveness within the field of tourism as a norm for evaluation. The study provides valuable perspectives on the domain-specific capabilities of artificial intelligence and offers indications of its wider potential and limitations within several academic disciplines. Furthermore, the findings of the study about the competence of AI in producing and addressing questions have the potential to bring about an important change in the design of assessments. In future examinations, there is a potential shift towards emphasizing the assessment of advanced intellectual skills, such as critical thinking, innovation, and applied competencies, as opposed to simply assessing the ability to memorize and recollect information.

## 7. Recommendations

ChatGPT-4 has demonstrated considerable achievement in fulfilling the criteria of intellectual thinking standards and exhibiting writing talents that closely resemble those of humans. The circumstance mentioned above presents a potential challenge to administering examinations in the context of tourism education. This section presents a set of recommendations for educators to reduce this threat.

Coghlan et al. (2020) highlight that implementing online examination supervision systems has effectively reduced cases of academic dishonesty and unethical behaviour. These systems specifically address behaviours such as copying test answers from peers or online sources and the unauthorized use of cheating technologies. Certain universities have offered reasons for utilizing these systems, whereas others have decided to cease or decline their implementation. The author warns that despite the implementation of these regulations, there is a possibility of certain students engaging in academic dishonesty.

There is a diverse range of regulations that can be employed to detect and reduce cases of academic dishonesty. The regulations include various measures, such as identity verification, keyboard recognition, video proctoring, browser control, test duration limitations, question and answer randomization, the creation of cheat-resistant questions, and reducing the significance of the exam in the overall course marking. Offline or in-person proctored examinations may present a more effective obstacle against cheating than online exams (Noorbehbahani et al., 2022).

The findings of a study indicate that the utilization of online proctoring has the potential to serve as an efficacious strategy in addressing instances of academic dishonesty within the context of online courses. Nevertheless, the research also underscores the necessity for continuous improvement in all aspects of distance education, including the level of instruction and knowledge acquisition, active participation, fairness, inclusivity, and moral behaviour. Hence, higher education institutions that provide online courses must consider the incorporation of online proctoring as a fundamental

component of a holistic strategy aimed at maintaining academic and exam integrity within online learning settings (Dendir & Maxwell, 2020). Potential solutions for this problem include advancing innovative plagiarism detection mechanisms with enhanced capabilities in distinguishing AI-generated content (Skavronskaya et al., 2023).

In general, various methods can be implemented to prevent the utilization of ChatGPT-4 or similar tools for academic dishonesty in examinations or assignments, particularly in the domain of tourism education. The use of online exam supervision or proctoring tools may reduce academic dishonesty. During exams, these systems can monitor students in real time and look for suspicious behaviour, like moving away from the exam tab or using another device. Using identity verification makes sure that the student who signed up for the test is the one who is taking it. Keyboard recognition can determine if someone is using external tools or software by looking for trends in how they write. The advancement of new plagiarism detection techniques capable of discerning AI-generated text is crucial to upholding online exam and assignment integrity. Offline or in-person supervised examinations can more efficiently reduce cheating than the web methods. The use of a restrictive time limit during examinations can effectively limit the opportunity for potential dishonesty. The use of randomization in the assignment of questions ensures that every student is presented with an individual set of questions, hence increasing the difficulty of sharing answers. Designing questions that require the utilization of critical thinking, practical application, and analytical skills may deter students from seeking direct responses from AI systems. In conclusion, a combination of technology solutions, improvements to exam design, and a focus on ethical behaviour can achieve the effective prevention of risks associated with tools like ChatGPT-4 in exams and assignments.

Studies should be conducted across various academic disciplines to investigate the critical thinking capabilities and human-like writing abilities of ChatGPT-4. It is essential to determine the abilities or limitations of this chatbot in many academic fields and to conduct a comprehensive assessment covering various question formats, such as multiple-choice and open-ended questions. This evaluation will enable an in-depth evaluation of its performance in an examination setting.

## Limitations and directions for further studies

The primary emphasis of this study focuses on the performance evaluation of ChatGPT-4 concerning intellectual criteria. The model's range may not be comprehensive regarding its possible uses or potential misapplications across various academic and non-academic contexts. The findings and recommendations could apply particularly to the context of tourism education. It is uncertain whether these findings may be readily generalised to other disciplines or fields of study without conducting more studies. The discussion seems to be centred around



hypothetical scenarios and evaluations rather than empirical studies that involve actual students, educators, or examination settings. The methodology, technologies, and issues associated with online education are subject to continual change. The material may not completely include all the details or predict future difficulties within online examinations.

Further studies should be conducted across various academic disciplines to investigate the critical thinking capabilities and human-like writing abilities of ChatGPT-4. It is essential to determine the abilities or limitations of this chatbot in many academic fields. Furthermore, it is essential to subject ChatGPT-4 to a comprehensive assessment covering various question formats, such as multiple-choice and open-ended questions. This evaluation will enable an in-depth evaluation of its performance in an examination setting.

**Declaration of competing interest:** None

## References

- Ababneh, K. I., Ahmed, K., & Dedousis, E. (2022). Predictors of cheating in online exams among business students during the Covid pandemic: Testing the theory of planned behavior. *The International Journal of Management Education*, 20(3), 100713. <https://doi.org/10.1016/j.ijme.2022.100713>
- Adzima, K. (2021). Examining online cheating in higher education using traditional classroom cheating as a guide. *Electronic Journal of E-Learning*, 18(6). <https://doi.org/10.34190/JEL.18.6.002>
- Agasisti, T., & Soncin, M. (2021). Higher education in troubled times: On the impact of Covid-19 in Italy. *Studies in Higher Education*, 46(1), 86–95. <https://doi.org/10.1080/03075079.2020.1859689>
- AlAfnan, M. A., Dishari, S., Jovic, M. & Lomidze, K. (2023). Chatgpt as an educational tool: opportunities, challenges, and recommendations for communication, business writing, and composition courses. *Journal of Artificial Intelligence and Technology*, 3, 60-68. <https://doi.org/10.37965/jait.2023.0184>
- Ali, K., Barhom, N., Marino, F. T., & Duggal, M. (2023). *The thrills and chills of chatgpt: implications for assessments in undergraduate dental education* [Preprint]. MEDICINE & PHARMACOLOGY. <https://doi.org/10.20944/preprints202302.0513.v1>
- Atlas, S. (2023). *Chatgpt for higher education and professional development: a guide to conversational AI*. University of Rhode Island DigitalCommons@URI. [https://digitalcommons.uri.edu/cba\\_facpubs/548](https://digitalcommons.uri.edu/cba_facpubs/548)
- Barrientos, A., Cuadros, M., Alba, J., & Cruz, A. S. (2021, October 27-29). *Implementation of a remote system for the supervision of online exams through the use of cameras with artificial intelligence*. 2021 IEEE Engineering International Research Conference (EIRCON), 1–4. <https://doi.org/10.1109/EIRCON52903.2021.9613352>
- Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, 57, 100745. <https://doi.org/10.1016/j.asw.2023.100745>
- Benvenuti, M., Cangelosi, A., Weinberger, A., Mazzoni, E., Benassi, M., Barbaresi, M., & Orsoni, M. (2023). Artificial intelligence and human behavioral development: A perspective on new skills and competences acquisition for the educational context. *Computers in Human Behavior*, 107903. <https://doi.org/10.1016/j.chb.2023.107903>
- Bernabei, M., Colabianchi, S., Falegnami, A., & Costantino, F. (2023). Students' use of large language models in engineering education: A case study on technology acceptance, perceptions, efficacy, and detection chances. *Computers and Education: Artificial Intelligence*, 100172. <https://doi.org/10.1016/j.caeai.2023.100172>
- Carvalho, L., Martinez-Maldonado, R., Tsai, Y.-S., Markauskaite, L., & De Laat, M. (2022). How can we design for learning in an AI world? *Computers and Education: Artificial Intelligence*, 3, 100053. <https://doi.org/10.1016/j.caeai.2022.100053>
- Chirumamilla, A., & Sindre, G. (2019). *Mitigation of cheating in online exams: Strengths and limitations* (A. V. S. Kumar, Ed.). IGI Global. <https://doi.org/10.4018/978-1-5225-7724-9>
- Choi, J. H., Hickman, K. E., Monahan, A., & Schwarcz, D. B. (2023). ChatGPT goes to law school. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4335905>
- Coghlan, S., Miller, T., & Paterson, J. (2020). Good proctor or “Big Brother”? AI ethics and online exam supervision technologies. <https://arxiv.org/abs/2011.07647>.
- Comas-Forgas, R., Lancaster, T., Calvo-Sastre, A., & Sureda-Negre, J. (2021). Exam cheating and academic integrity breaches during the COVID-19 pandemic: An analysis of internet search activity in Spain. *Heliyon*, 7(10), e08233. <https://doi.org/10.1016/j.heliyon.2021.e08233>
- 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>
- Crawford, J., Cowling, M., Central Queensland University, Australia, Allen, K.-A., & Monash University, Australia. (2023). Leadership is needed for ethical ChatGPT: Character, assessment, and learning using artificial intelligence (AI). *Journal of University Teaching and Learning Practice*, 20(3). <https://doi.org/10.53761/1.20.3.02>
- Currie, G. M. (2023). Academic integrity and artificial intelligence: Is ChatGPT hype, hero or heresy? *Seminars in Nuclear Medicine*, 53(5), 719–730. <https://doi.org/10.1053/j.semnucmed.2023.04.008>
- David, A., Mihai, D., Mihailescu, M.-E., Carabas, M., & Tapus, N. (2022). Scalability through distributed deployment for moodle learning management system. *Procedia Computer Science*, 214, 34–41. <https://doi.org/10.1016/j.procs.2022.11.145>
- Dendir, S., & Maxwell, R. S. (2020). Cheating in online courses: Evidence from online proctoring. *Computers in Human Behavior Reports*, 2, 100033. <https://doi.org/10.1016/j.chbr.2020.100033>
- Dias, J., Kamdi, D., Gharat, N., & Chudhari, P. (2019). Chatbot for Government Examination using AI. *Journal of Engineering*, 9(4), 58-62.
- Duhaim, A. M., Al-mamory, S. O., & Mahdi, M. S. (2022). Cheating detection in online exams during Covid-19 pandemic using data mining techniques. *Webology*, 19(1), 341–366. <https://doi.org/10.14704/WEB/V19I1/WEB19026>
- Eke, D. O. (2023). ChatGPT and the rise of generative AI: Threat to academic integrity? *Journal of Responsible Technology*, 13, 100060. <https://doi.org/10.1016/j.jrt.2023.100060>
- Elbanna, S., & Armstrong, L. (2023). Exploring the integration of ChatGPT in education: Adapting for the future. *Management & Sustainability: An Arab Review*. <https://doi.org/10.1108/MSAR-03-2023-0016>
- Essel, H. B., Vlachopoulos, D., Tachie-Menson, A., Johnson, E. E., & Baah, P. K. (2022). The impact of a virtual teaching assistant (chatbot) on students' learning in Ghanaian higher education. *International Journal of Educational Technology in Higher Education*, 19(1), 57. <https://doi.org/10.1186/s41239-022-00362-6>
- Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2023). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*, 1–15. <https://doi.org/10.1080/14703297.2023.2195846>
- Geerling, W., Mateer, G. D., Wooten, J., & Damodaran, N. (2023a). ChatGPT has Mastered the Principles of Economics: Now What? *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4356034>
- Gerashchenko, I. P., & Kovalev, V. A. (2021). *Formation of educational ecosystems through the digital transformation of the educational environment*. SHS Web of Conferences, 121, 03004. <https://doi.org/10.1051/shsconf/202112103004>





- Gilson, A., Safranek, C., Huang, T., Socrates, V., Chi, L., Taylor, R. A., & Chartash, D. (2022). *How Does ChatGPT Perform on the Medical Licensing Exams? The implications of large language models for medical education and knowledge assessment* [Preprint]. Medical Education. <https://doi.org/10.1101/2022.12.23.22283901>
- Göktaş, L. S. (2023). ChatGPT uzaktan eğitim sınavlarında başarılı olabilir mi turizm alanında doğruluk ve doğrulama üzerine bir araştırma (Can ChatGPT succeed in distance education exams a research on accuracy and verification in tourism). *Journal of Tourism and Gastronomy Studies*, 2. <https://doi.org/10.21325/jotags.2023.1224>
- Hashim, H. (2018). Application of technology in the digital era education. *International Journal of Research in Counseling and Education*, 1(2). <https://doi.org/10.24036/002za0002>
- Hill, B. (2023). *Taking the help or going alone: ChatGPT and class assignments*. SSRN Electronic Journal. <https://doi.org/10.2139/ssrn.4465833>
- Holden, O. L., Norris, M. E., & Kuhlmeier, V. A. (2021). Academic integrity in online assessment: A research review. *Frontiers in Education*, 6, 639814. <https://doi.org/10.3389/educ.2021.639814>
- Hwang, A. (2018). Online and Hybrid Learning. *Journal of Management Education*, 42(4), 557–563. <https://doi.org/10.1177/1052562918777550>
- Ilynykh, S. (2021). *Online education: Trends and risks of digital technologies: VIII International Scientific and Practical Conference "Current problems of social and labour relations" (ISPC-CPSLR 2020)*, Makhachkala, Russia Federation. <https://doi.org/10.2991/assehr.k.210322.126>
- Jungherr, A. (2023). *Using ChatGPT and Other Large Language Model (LLM) applications for academic paper assignments* [Preprint]. SocArXiv. <https://doi.org/10.31235/osf.io/d84q6>
- King, D. L., & Case, C. J. (2014). E-Cheating: Incidence and trends among college students. *Issues In Information Systems*, 15(1). [https://doi.org/10.48009/1\\_iis\\_2014\\_20-27](https://doi.org/10.48009/1_iis_2014_20-27)
- Klijn, F., Mdaghri Alaoui, M., & Vorsatz, M. (2022). Academic integrity in on-line exams: Evidence from a randomized field experiment. *Journal of Economic Psychology*, 93, 102555. <https://doi.org/10.1016/j.joep.2022.102555>
- 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>
- Lanier, M. M. (2006). Academic integrity and distance learning. *Journal of Criminal Justice Education*, 17(2), 244–261. <https://doi.org/10.1080/10511250600866166>
- Latifah, R., Budiyanto, C. W., & Saputro, H. (2022). Digital transformation readiness in education: A review. *International Journal of Information and Education Technology*, 12(8), 809–815. <https://doi.org/10.18178/ijiet.2022.12.8.1688>
- Lin, J. C., Younessi, D. N., Kurapati, S. S., Tang, O. Y., & Scott, I. U. (2023). Comparison of GPT-3.5, GPT-4, and human user performance on a practice ophthalmology written examination. *Eye*. <https://doi.org/10.1038/s41433-023-02564-2>
- Maslov, I., Nikou, S., & Hansen, P. (2021). Exploring user experience of learning management system. *The International Journal of Information and Learning Technology*, 38(4), 344–363. <https://doi.org/10.1108/IJILT-03-2021-0046>
- Mellieon, Jr, H. I., & Robinson, P. A. (2021). The New norm: Faculty Perceptions of condensed online learning. *American Journal of Distance Education*, 35(3), 170–183. <https://doi.org/10.1080/08923647.2020.1847626>
- Misra, D. P., & Chandwar, K. (2023). ChatGPT, artificial intelligence and scientific writing: What authors, peer reviewers and editors should know. *Journal of the Royal College of Physicians of Edinburgh*, 53(2), 90–93. <https://doi.org/10.1177/14782715231181023>
- Montenegro-Rueda, M., Luque-de La Rosa, A., Sarasola Sánchez-Serrano, J. L., & Fernández-Cerero, J. (2021). Assessment in higher education during the Covid-19 pandemic: A systematic review. *Sustainability*, 13(19), 10509. <https://doi.org/10.3390/su131910509>
- Naidu, K., & Sevnarayan, K. (2023). ChatGPT: An ever-increasing encroachment of artificial intelligence in online assessment in distance education. *Online Journal of Communication and Media Technologies*, 13(3), e202336. <https://doi.org/10.30935/ojcm/13291>
- Newton, P. M., & Xiromeriti, M. (2023). *ChatGPT Performance on MCQ Exams in Higher Education. A Pragmatic Scoping Review* [Preprint]. <https://doi.org/10.35542/osf.io/sytu3>
- Nguyen, X.-A., Nguyen, S.-N., Nguyen, T.-T.-T., Luong, D.-H., & Tran, T.-G. (2022). Exploring Some academic dishonesty in remote online exams of vietnamese high school students in context of the Covid-19 pandemic. *Asian Journal of Education and Social Studies*, 13–26. <https://doi.org/10.9734/ajess/2022/v33i130782>
- Nieuwoudt, J. E. (2020). Investigating synchronous and asynchronous class attendance as predictors of academic success in online education. *Australasian Journal of Educational Technology*, 36(3), 15–25. <https://doi.org/10.14742/ajet.5137>
- Noorbehbahani, F., Mohammadi, A., & Aminazadeh, M. (2022). A systematic review of research on cheating in online exams from 2010 to 2021. *Education and Information Technologies*, 27(6), 8413–8460. <https://doi.org/10.1007/s10639-022-10927-7>
- Oravec, J. A. (2023). Artificial intelligence implications for academic cheating: expanding the dimensions of responsible human-AI collaboration with Chatgpt and Bard. *Jl. of Interactive Learning Research*, 34(2).
- Paul, R. (2005). The state of critical thinking today. *New Directions for Community Colleges*, 2005(130), 27–38. <https://doi.org/10.1002/cc.193>
- Paul, R., & Elder, L. (2013). Critical thinking: Intellectual standards essential to reasoning well within every domain of human thought, Part Two. *Journal of Developmental Education*, 37(1).
- Perera, P., & Lankathilaka, M. (2023). AI in higher education: A literature review of chatgpt and guidelines for responsible implementation. *International Journal of Research and Innovation in Social Science*, VII(VI), 306–314. <https://doi.org/10.47772/IJRISS.2023.7623>
- Pérez, J. Q., Daradoumis, T., & Puig, J. M. M. (2020). Rediscovering the use of chatbots in education: A systematic literature review. *Computer Applications in Engineering Education*, 28(6), 1549–1565. <https://doi.org/10.1002/cae.22326>
- Petrolo, D., Fakhari Manesh, M., & Palumbo, R. (2023). Unpacking business, management, and entrepreneurship education online: Insights from a hybrid literature review. *The International Journal of Management Education*, 21(2), 100812. <https://doi.org/10.1016/j.ijme.2023.100812>
- Putra, F. W., Rangka, I. B., Aminah, S., & Aditama, M. H. R. (2023). ChatGPT in the higher education environment: Perspectives from the theory of high order thinking skills. *Journal of Public Health*. <https://doi.org/10.1093/pubmed/fdad120>
- Rahm, L. (2023). Educational imaginaries: Governance at the intersection of technology and education. *Journal of Education Policy*, 38(1), 46–68. <https://doi.org/10.1080/02680939.2021.1970233>
- Skavronskaya, L., Hadinejad, A. (Hana), & Cotterell, D. (2023). Reversing the threat of artificial intelligence to opportunity: A discussion of ChatGPT in tourism education. *Journal of Teaching in Travel & Tourism*, 23(2), 253–258. <https://doi.org/10.1080/15313220.2023.2196658>
- Smutny, P., & Schreiberova, P. (2020). Chatbots for learning: A review of educational chatbots for the Facebook Messenger. *Computers & Education*, 151, 103862. <https://doi.org/10.1016/j.compedu.2020.103862>
- Susnjak, T. (2022). *ChatGPT: The End of Online Exam Integrity?*. <http://arxiv.org/abs/2212.09292>
- Swiecki, Z., Khosravi, H., Chen, G., Martinez-Maldonado, R., Lodge, J. M., Milligan, S., Selwyn, N., & Gašević, D. (2022). Assessment in the age of





artificial intelligence. *Computers and Education: Artificial Intelligence*, 3, 100075. <https://doi.org/10.1016/j.caeai.2022.100075>

Tsai, M.-L., Ong, C. W., & Chen, C.-L. (2023). Exploring the use of large language models (LLMs) in chemical engineering education: Building core course problem models with Chat-GPT. *Education for Chemical Engineers*, 44, 71–95. <https://doi.org/10.1016/j.ece.2023.05.001>

Ventayen, R. J. M. (2023). ChatGPT by OpenAI: Students' viewpoint on cheating using Artificial Intelligence-Based application. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4361548>

Von Grunigen, D., De Azevedo E Souza, F. B., Pradarelli, B., Magid, A., & Cieliebak, M. (2018). Best practices in e-assessments with a special focus on cheating prevention. *2018 IEEE Global Engineering Education Conference (EDUCON)*, 893–899. <https://doi.org/10.1109/EDUCON.2018.8363325>

Wagner, G., Lukyanenko, R., & Paré, G. (2022). Artificial intelligence and the conduct of literature reviews. *Journal of Information Technology*, 37(2), 209–226. <https://doi.org/10.1177/02683962211048201>

Watson, G. R., & Sottile, J. (2010). *Cheating in the digital age: do students cheat more in online courses?* [https://mds.marshall.edu/cgi/viewcontent.cgi?article=1000&context=eft\\_faculty](https://mds.marshall.edu/cgi/viewcontent.cgi?article=1000&context=eft_faculty)

Wiley. (2020). *White paper: Academic integrity in the age of online learning*. <https://www.wiley.com/en-us/network/education/instructors/teaching-strategies/academic-integrity-in-the-age-of-online-learning-3>

Xie, X., Siau, K., & Nah, F. F.-H. (2020). COVID-19 pandemic – online education in the new normal and the next normal. *Journal of Information Technology Case and Application Research*, 22(3), 175–187. <https://doi.org/10.1080/15228053.2020.1824884>

Yau, C., & Chan, K. (2023, February 17). *University of Hong Kong temporarily bans students from using ChatGPT, other AI-based tools for coursework*. <https://www.scmp.com/news/hong-kong/education/article/3210650/university-hong-kong-temporarily-bans-students-using-chatgpt-other-ai-based-tools-coursework>

Yorkovsky, Y., & Levenberg, I. (2022). Distance learning in science and mathematics—Advantages and disadvantages based on pre-service teachers' experience. *Teaching and Teacher Education*, 120, 103883. <https://doi.org/10.1016/j.tate.2022.103883>

Zaitsu, W., & Jin, M. (2023). Distinguishing ChatGPT(-3.5, -4)-generated and human-written papers through Japanese stylometric analysis. <https://arxiv.org/abs/2306.01761>