Jomat

Journal of multidisciplinary academic tourism
2024, 9 (3): 231-243
https://doi.org/10.31822/jomat.2024-9-3-231

ISSN: 2645-9078

The performance of artificial intelligence in the exams of tourist guidance

Abdullah Ülkü

Keywords:
Artificial intelligence,
Large language models (LLMs),
Tourism education,
Tourist guidance,
Exam.

Article History:
Submitted: 12.01.2024
Revised:17.05.2024
Revised:04.06.2023
Accepted: 27.06.2024
Published Online: 28.06.2024

ABSTRACT

The aim of this study is to evaluate the efficiency of ChatGPT versions 3.5 and 4 for training tourist guides. The study followed a systematic approach by conducting assessments on undergraduate students from three institutions who are enrolled in tourist guide education programs and both ChatGPT versions. Competent academicians assessed a succession of questions in the form of open-ended and multiple-choice questions. The mean scores obtained on the multiple-choice test for ChatGPT-4 were better than those of both ChatGPT-3.5 and college students, thereby indicating that ChatGPT-4 has greatly improved. Nevertheless, when responding to open-ended queries, individuals with real-life experience as tour guides gave much more inclusive as well as convincing answers compared to ChatGPT-4. This underscores the importance of hands-on experiences in training tour guides, where AI technology is currently weak. This study contributes to better comprehension regarding the role played by artificial intelligence (AI) in education with reference to the tourism industry specifically. While at the same time emphasizing how critical human expertise is needed during practical learning sessions, this implies that AI has potential for disseminating theoretical knowledge. The results suggest that AI is a beneficial supplementary aid in educational environments, rather than a replacement for human-centered instructional approaches.

1. Introduction

Artificial Intelligence (AI) is transforming various sectors, including the field of education. Integrating AI, especially chatbots, into educational environments has the capacity to revolutionize conventional teaching techniques, improve learning experiences, and offer individualized educational assistance to students. A chatbot "*is a software tool that interacts with users on a certain topic or in a specific domain in a natural, conversational way using text and voice*." (Smutny & Schreiberova, 2020). Chatbots have been used across a wide range of domains, including education (Essel et al., 2022; Okonkwo & Ade-Ibijola, 2021); marketing (Lin et al., 2022; Jenneboer et al., 2022; Luo et al., 2023; Xu et al., 2023); e-commerce (Wang et al., 2023; Li & Wang, 2023); tourism (Leung & Wen, 2020; Melián-González et al., 2021; Orden-Mejía & Huertas, 2022); and medicine (Athavale et al., 2023; L. Li, 2023).

One of the most famous chatbots is ChatGPT, released by OpenAI on November 30, 2022. It has attracted millions of users pretty quickly (Hartmann et al., 2023). ChatGPT is an advanced natural language processing model. Aydin & Karaarslan (2023) pre-trained the model on a large dataset, often making it indistinguishable from human-written content. Since its release, ChatGPT has been the subject of numerous academic studies in various fields such as education (Currie et al., 2023; Keiper et al., 2023); academic writing (Dalalah & Dalalah, 2023); argumentative writing (Su et al., 2023); medical writing (Ho et al., 2023); medicine (Jin & Dobry, 2023; W. Li et al., 2023; Ma, 2023); entrepreneurship (Short & Short, 2023); language learning (Kohnke et al., 2023); philosophy (Peters et al., 2023); and finance (Cao & Zhai, 2023). Though ChatGPT has been the subject of research in a number of disciplines, none yet have focused on tourist guides.

Tourist guides are employees who directly and indirectly influence service quality (Tsaur & Teng, 2017), tourists' experiences, and satisfaction, playing a significant role in the tourism industry (Alazaizeh et al., 2019). They are guides, communicators, and experience brokers (Parsons et al., 2019). They interpret the cultural and natural heritage of an area. They have multiple roles (care role, instrumental role, social role, interactional role, communicative role, and dealing with emergency role). (Tsaur & Teng, 2017). They have communicative skills and destination knowledge (Zhu & Xu, 2021). Given the multifaceted responsibilities of tourist guides, their training process is crucial and tourist guides should receive proper

Research Paper

Abdullah Ülkü: *Assit. Prof., Harran University, Şanlıurfa, Türkiye, Email: abdullahulku@harran.edu.tr, Orcid Id: 0000-0002-0937-2252*

training in order to fulfill all their roles and responsibilities effectively (Tsaur & Teng, 2017).

Tourist guide training is important for a country's tourism industry. This study aimed to evaluate the precision and dependability of the data supplied by ChatGPT during this important training process. In this regard, when we have examined international databases like Web of Science, EBSCOhost, ProQuest, SSCI, and Scopus, etc., we have used "AI in education, AI and tourism, AI and tourist guide, ChatGPT in education, ChatGPT exam, ChatGPT performance, ChatGPT in tourism, ChatGPT in tourism education, ChatGPT and tourist guide" words. As a result, we have not come across a study that evaluates the exam performance of ChatGPT in the field of tourist guide.

ChatGPT might be able to succeed in university exams (Leiter et al., 2023), and there are studies available that indicate the feasibility of this (Bommarito II & Katz, 2022; Choi et al., 2023; Kung et al., 2023; J. C. Lin et al., 2023; Nunes et al., 2023). ChatGPT obtains information from various sources, including books, articles, websites, and other online content. It uses this corpus to generate appropriate responses. It has access to current information, and the corpus undergoes continuous updates to maintain accuracy. Additionally, this chatbot does not forget previous conversations and corrects itself if errors occur. Released in March 2023, GPT-4 can accept visual queries, handle 32,000 tokens simultaneously, and has 100 trillion parameters (Barrot, 2023) whereas chatGPT-3.5 has 175 billion (Zaitsu & Jin, 2023). ChatGPT-3.5's knowledge spans from September 2021 and is available for free use.

There exists a significant research gap concerning the understanding of the specific impacts and effectiveness of advanced AI models, such as ChatGPT, in the context of taking tourist guide exams and their training processes. While much research has been done on the use of AI in the many scopes related to education, none are focused on trying to evaluate the performance of AI in tourist guide training and exam performance. This will help fill the existing gap and examine the extent to which ChatGPT has performed when used for tourist guides' training.

This research adds to the academic knowledge about AI in education, specifically in the context of tourism and training for tourist guides. It evaluates the performance of ChatGPT in a particular domain, offering insights into the customization of AI for different fields of study. The tourism industry recognizes the immense potential of AI in improving the training of tourist guides, who play a vital role in shaping tourists' experiences and overall satisfaction. This study provides insights into the abilities and constraints of AI in an academic environment by comparing ChatGPT with undergraduate students. Overall, it contributes to the academic discussion on utilizing AI in tourist guide exams and the potential advantages for their training.

In our study, we conducted exams for ChatGPT (3.5 and 4) and undergraduate students studying in the tourist guidance department of three universities. Questions on the exam belong to critical lessons for tourist guide training. Expert academics have reviewed all the questions. This study aimed to determine whether ChatGPT (3.5 and 4) outperforms students and to evaluate the precision and dependability of the data it provides. In this regard, the research questions for this study are below:

RQ1: Can ChatGPT 3.5 be more successful than students in the tourist guide exam?
RQ2: Can ChatGPT 4 be more successful than students in the tourist guide exam?
RQ3: Are ChatGPT 3.5's answers reliable in the tourist guide exam?
RQ4: Are ChatGPT 4's answers reliable in the tourist guide exam?

The findings of this study indicate that ChatGPT-4 performed better than both ChatGPT-3.5 and the students in multiple-choice questions, showing considerable progress in AI technology. On the contrary, on open-ended questions, actual experienced students were able to give more detailed as well as convincing answers than ChatGPT-4. Therefore, these findings confirm that it is fundamental as a tourist guide to have practical work experience which AI is currently not equipped with for reproducing such kind of knowledge.

2. Literature Review

AI in Education

AI encompasses a science called artificial intelligence, which is used to simulate human behavior on computers and trains computers to learn human behaviors (Zhang & Lu, 2021). Artificial intelligence is composed of many disciplines and has made significant advances in areas such as language learning (Pikhart, 2020); speech recognition (Iqbal & Faiz, 2020); healthcare industry (Lee & Yoon, 2021); energy sustainability (Chui et al., 2018); transport (Abduljabbar et al., 2019); building (Dinesh et al., 2023); agrifood sectors (Kutyauripo et al., 2023); economy (F. Wang et al., 2023); education (Chen et al., 2020); and tourism (Doborjeh et al., 2022).

There has been a rise in the utilization of AI for educational purposes recently (Gimpel et al., 2023). Instead of following conventional methods, AI applications have surpassed the conventional understanding in the educational field. AI, often referred to as a high-capacity computer, encompasses traditional computer systems as well. AI has the capability to enrich students' learning experiences starting from early childhood with the assistance of tools supporting artificial intelligence and computers (Chen et al., 2020). AI is positioned to assist students in exam preparation and lesson comprehension (Pursnani et al., 2023). It has the capability to impact courses across various fields. AI can be utilized by students to address exam queries and complete their assignments

(Chassignol et al., 2018). Moreover, AI can offer tailored assistance to students and aid in the enhancement of their communication and writing proficiencies (Tyson & Sauers, 2021).

Students can use AI for language generation and processing (Atlas, 2023), information counselling (Y. Lin & Yu, 2023), grammar assistance, and answering questions (Rudolph et al., 2023). AI has offered opportunities for educators. They can develop their knowledge and flexible skills for a constantly changing world (Chen et al., 2020), improve student outcomes by gaining the data to understand students' needs and personalize learning activities (Chassignol et al., 2018), provide feedback more easily to students, cope with an increased workload, and tailor their instruction to the needs of their students (Singh & Hiran, 2022). Educators can evaluate information and process it faster thanks to AI. Also, classroom management tasks can become easier, and assessment of learning can be facilitated, such as grading student essays (Tyson & Sauers, 2021).

One of the opportunities AI offers to education system is chatbots. A chatbot is a tool that combines language processing or artificial intelligence and interacts with people through text or voice (Pérez et al., 2020). Chatbots have the ability to discern patterns from input data and then generate corresponding results based on the provided input. (Agarwal et al., 2022). A chatbot system can enhance students' academic performance, critical thinking, and learning satisfaction (Chang et al., 2022).

Chatbots are research assistants and can be used for exams and assessments (Kooli, 2023). Educators can use chatbots to create exam questions. Educators and students can ask chatbots any exam question, and chatbots can answer that question (Dias et al., 2019); but here a problem arises. Are the answers given by chatbots accurate and reliable? This research aims to assess the performance and dependability of ChatGPT, a renowned chatbot well recognized globally. The following section provides an overview of the examinations in which ChatGPT has undergone testing, along with the corresponding outcomes achieved.

ChatGPT in exams

ChatGPT is an artificial intelligence chatbot developed by OpenAI. OpenAI is an AI research and deployment company, and their mission is to ensure that artificial general intelligence benefits all of humanity (OpenAIb, n.d.). GPT-3.5 and GPT-4 can contextualize text information and generate appropriate responses (Susnjak, 2022). GPT-3.5 accepts text input and produces text output. Beyond that, GPT-4 also accepts images as part of the input prompt (OpenAI, n.d.). GPT-4 can accept visual queries, handle 32,000 tokens simultaneously, and has 100 trillion parameters (Barrot, 2023) whereas chatGPT-3.5 has 175 billion (Zaitsu & Jin, 2023). The knowledge possessed by ChatGPT-3.5 is limited to September 2021 and is free to use.

ChatGPT, particularly GPT-4, exhibited a notable degree of accuracy and ability in various domains such as law (Bommarito II and Katz, 2022), medicine (J. C. Lin et al., 2023), and economics (Geerling et al., 2023). This trend extends to economics and social sciences, where ChatGPT answered a significant portion of questions correctly, often outperforming students.

Studies have evaluated ChatGPT's performance across various fields, highlighting its potential and limitations. In legal education, Bommarito II and Katz (2022) found GPT-3.5 scored 50.3% on the Bar Examination, exceeding expectations, while Choi et al. (2023) showed it achieved satisfactory grades in four law school finals at the University of Minnesota. In the medical field, Gilson et al. (2022) reported that GPT-3.5 scored over 60% on the USMLE Step-1, similar to a third-year medical student. Kung et al. (2023) confirmed these results, and J. C. Lin et al. (2023) found that GPT-4 outperformed both residents and practicing ophthalmologists on a 260-question ophthalmology exam. In economics, Geerling et al. (2023) demonstrated that ChatGPT correctly answered 63.3% of microeconomics and 86.7% of macroeconomics questions, outperforming many students. They recommended emphasizing experiential learning and limiting unsupervised assessments to prevent overreliance on AI.

Despite its remarkable capabilities, ChatGPT exhibits deficiencies in assessments that require practical knowledge, critical thinking, and complex problem-solving. ChatGPT's performance has demonstrated variability in disciplines that require a comprehensive understanding of practical applications and analytical reasoning, such as physics and complex engineering tasks, which has raised concerns about its dependability. Despite the fact that ChatGPT offers pertinent insights, research on life support evaluations (Fijačko et al., 2023), physics (Kortemeyer, 2023), and specific engineering scenarios (Naser et al., 2023) indicates that it fails to consistently meet the anticipated benchmarks. This emphasizes its function as a supplementary learning device rather than a standalone solution.

ChatGPT was ineffectual in its evaluation of life support examinations by Fijačko et al. (2023), despite the fact that they provided relevant and accurate explanations. This emphasizes the model's function as a study aid rather than a definitive answer provider. Nevertheless, it failed to perform well in physics duties. In a study conducted by Kortemeyer (2023), it was shown that ChatGPT was unable to obtain the minimal passing score for an introductory physics course. This raises concerns regarding the integration of AI into education, particularly in light of the fact that the majority of graduates are anticipated to work with AI in their future professions. Similarly, Naser et al. (2023) investigated civil and environmental engineering in a different engineering discipline. They contrasted the responses provided by ChatGPT with those recommended by the National Council of Examiners for

Engineering and Surveying. ChatGPT-4's efficacy in the engineering domain is inconsistent, as evidenced by its score of 70.9% on one exam and its score of 46.2% on another engineering-context exam.

When it comes to undergraduate education, ChatGPT has demonstrated its proficiency in addressing intricate subjects. For instance, it achieved a score of 20.5 out of 40 in the algorithms and data structures section, which suggests that it has a comprehensive understanding of the subject matter (Bordt & von Luxburg, 2023). Similarly, GPT-4 outperformed GPT-3.5 by 11 points on a multidisciplinary Brazilian university entrance exam, achieving an 87% accuracy rate (Nunes et al., 2023).

ChatGPT has also demonstrated efficacy in the health sciences, particularly in the field of dental education. It excelled in a variety of examinations, providing precise responses and producing valuable reflection reports and research evaluations (Ali et al., 2023). Giannos and Delardas (2023) emphasized the strengths of ChatGPT in critical thinking, textual comprehension, and problem-solving. Nevertheless, it encountered challenges with intricate mathematical reasoning. Victor et al. (2023) evaluated ChatGPT in the field of social work by administering examinations from the Association of Social Work Boards in the United States. The model achieved a score of 76%, 80%, and 64% on the Bachelor's, Master's, and Clinical exams, respectively, indicating a high level of success in the field of social work reasoning.

Newton & Xiromeriti (2023) analyzed 41 studies in which ChatGPT was tested with a total of 46204 multiple-choice questions (MCQs) from 97 question sets. They found that versions of ChatGPT based on GPT-4 passed most MCQ-based exams and got similar results with most human participants. Versions of ChatGPT based on GPT-3 or 3.5 performed worse than the average human student. These 41 studies are about medicine, dentistry, computer science, life support, business, law, economics, math, parasitology, physics, chemistry, social work, engineering, anatomy, and accounting. The author of this study examined 51 studies on the exam performance of ChatGPT. It is found that most of the studies (42) have been conducted in the field of science, and the author thinks that there are not enough studies in the field of social sciences. None of the studies in the social sciences are related to tourist guides.

ChatGPT's performance exhibits significant differences across many fields. Although it demonstrates proficiency in standardized assessments in the domains of law, medicine, and economics, it has difficulties when confronted with assignments in disciplines such as physics and complex engineering. The diversity emphasizes the necessity of making discipline-specific modifications and improvements to AI training data and algorithms. The different findings suggest that the adoption of ChatGPT in educational and professional environments will require careful adjustment in order to effectively harness its advantages while addressing its limitations. Figure 1 shows general findings about the performance of ChatGPT in exams.

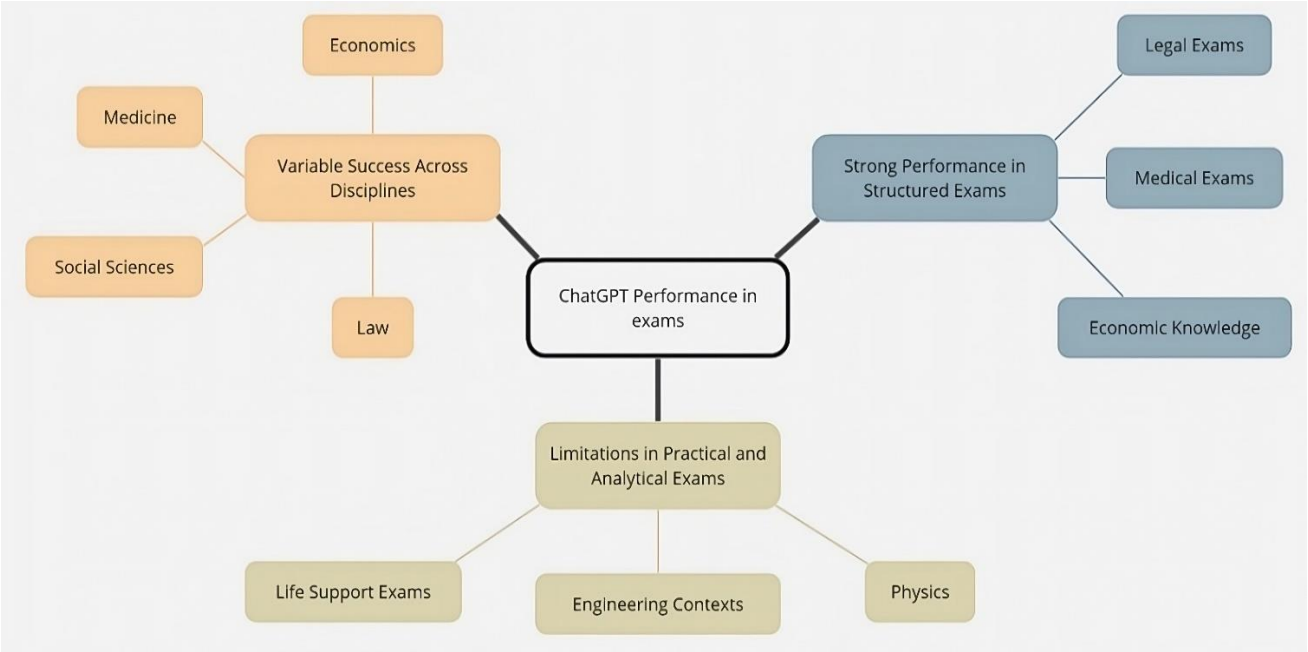

Figure 1. ChatGPT in exams

Source: Elaborated by Author

The results of some studies show that ChatGPT has the potential to serve as a useful tool in educational environments, specifically for the distribution of theoretical information and helping in exam preparing. The efficiency with which it can handle large amounts of organized information makes it a powerful tool for students and educators. Nevertheless, the limitations of AI in both practical and analytical fields suggest that it should serve as a supplement to human instruction rather than an alternative. In tourism and tourist guidance field, where hands-on experience and cultural sensitivity are of utmost importance, the human element continues to play a significant role in teaching practical skills and critical thinking.

3. Methodology

We conducted an exam for ChatGPT (3.5 and 4) and 124 undergraduate students studying in the tourist guidance department of three universities in the southeast region of Türkiye. The exam papers were sent to the lecturers working at the universities and the students took the exam in the designated classes. Questions were formatted into two variants and input into ChatGPT as follows:

1) Open-ended prompting: Created by removing all answer choices. A scenario was presented, and an answer was requested according to this scenario. For example, "According to the scenario above, what do travel agencies consider when selecting tourist guides?"

2) Multiple-choice single answer: ChatGPT and students could only choose one response from a predetermined list we defined. For example, "In which of the following cities was the first urban underground subway built?"

The exam consisted of 25 multiple-choice questions (MCQ) sourced from the website of the Association of Turkish Tourist Guides (TUREB, n.d.-b, n.d.-a) and other 25 MCQ prepared by expert academics. 25 questions from the Association of Turkish Tourist Guides (ATTG) belong to guidance certificate programs for two regions in Turkey. Questions directly related to a city in these regions have been removed, and general culture-level questions have been used. Answers have been evaluated according to the answer key for this association. Each of the 50 questions is worth two points.

The other 25 questions belong to critical lessons (travel agency and tour operator management, ancient civilizations of Anatolia, mythology, history of religions, and archaeology) for the tourist guidance training and have been prepared by authors. The students taking the exam were selected from those who had taken these lessons before. All questions have been reviewed by expert academics, and recommended corrections have been made. This article was approved by the ethics committee of Harran University. Number of ethics committee approval report: No. 2023/151.

Table 1. Lessons and number of questions

Field	N
General culture (ATTG)	25
Travel Agency and Tour Operator Management	5
Ancient Civilizations of Anatolia	5
Mythology	5
History of Religions	5
Archeology	5

Source: Elaborated by Authors

We have also used four open-ended questions. We have presented two scenarios to ChatGPT and asked it to answer the questions. The evaluation of open-ended questions was blind. Five faculty members working in the department of tourist guidance were asked to evaluate the answers given by the students. Two faculty members also evaluated the answers given by ChatGPT. ChatGPT gave wrong answers to some of the questions in Turkish. When the same questions were asked in English, it gave the correct answers. Therefore, all questions were asked in English.

4. Results

The exam of general culture (MCQ)

The data related to the general culture exam are presented in Table 2. This table compares the average scores of both ChatGPT versions and undergraduate students from three different universities. In this exam, ChatGPT-4 answered 24 out of 25 questions correctly, scoring 48 points and ranking first. ChatGPT-3.5 scored 36 points with 18 correct answers. This indicates an improvement in the newer version of ChatGPT over the previous one, highlighting significant advancements in artificial intelligence technology and an increased capacity to process a range of complex data.

Table 2. The results of the general culture exam

Participants	Scores	Number of correct answers
GPT-3.5	36	18
GPT-4	48	24
X university students	38.6	19.3
Y university students	34.8	17.4
Z university students	23.4	16.7

Source: Elaborated by Author. The scores and correct answers of the students are average results

It is observed that students from X University, with an average of 19.3 correct answers, performed better than ChatGPT-3.5. Students from the other universities, however, scored lower than both versions of ChatGPT. The findings provide insights into the capabilities of artificial intelligence-driven technologies in educational environments, particularly in the area of general cultural knowledge.

The exam of other lessons (MCQ)

The exam results for some courses that are important in tourist guide training are given in Table 3. Within the context of specialized academic fields, the findings shed light on the possibilities of artificial intelligence as well as

the limitations of this technology. The performance of both versions of ChatGPT has been satisfactory across all exams. One thing that is particularly interesting is the fact that ChatGPT-4 achieved the top marks in every field. The consistent performance of the AI models in these specialized subjects provides evidence of their significant breadth of knowledge and ability to access relevant data across various fields.

Table 3. The results of the other lessons exam

Lessons	GPT-3.5	GPT-4	X university students	Y university students	Z university students
Travel Agency and Tour Operator Management	3	5	4.2	4.1	3.5
Mythology	5	5	4.4	3.1	2.8
Ancient civilizations of Anatolia	4	5	4.1	3.7	3.1
History of religions	5	5	4.6	4.2	3.6
Archaeology	4	4	4.3	3.5	3.2
Total point	*42*	*48*	*43.2*	*37.2*	*32.4*

Source: Elaborated by Authors. The points and correct answers of the students are average results

The performance of university students exhibits variations across different subjects, with students from X University achieving a total score higher than that of ChatGPT-3.5. Students from Y University also scored closely to ChatGPT-3.5. These findings demonstrate the discrepancies in scores among the universities and emphasize the variability inherent in human learning processes. In Figure 1, the scores for each lesson are presented in color.

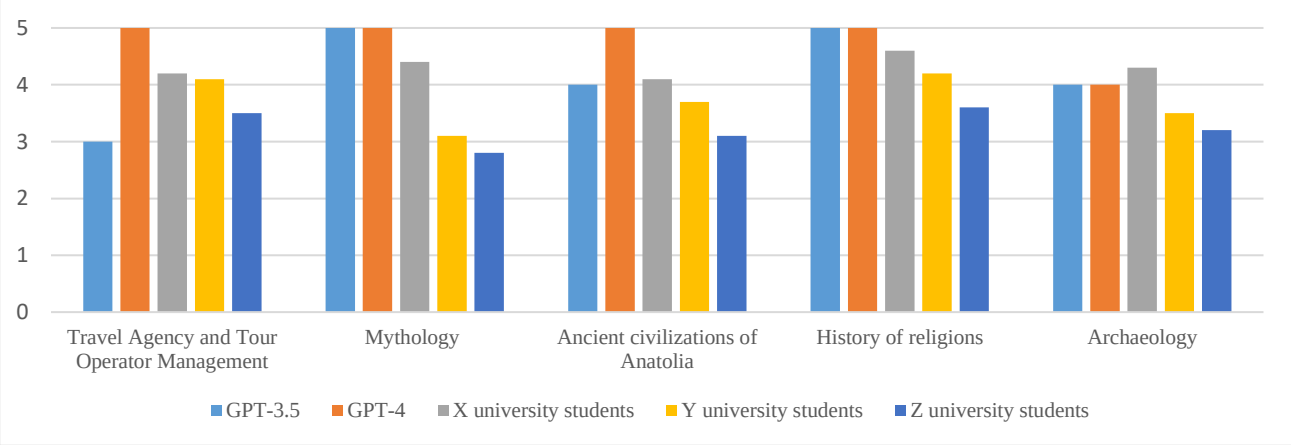

Figure. 2 The results of the other lessons exam

Source: Elaborated by Authors

ChatGPT-4 has achieved success in all areas, scoring 96 out of 100 in the multiple-choice examination. The total score obtained by students from X University has surpassed that of ChatGPT-3.5. Figure 2 presents the total scores obtained from all exams. While interpreting Figure 2. which displays the total scores of the MCQ exams, we observe a comparative performance analysis between AI models (specifically ChatGPT) and learners across various academic disciplines.

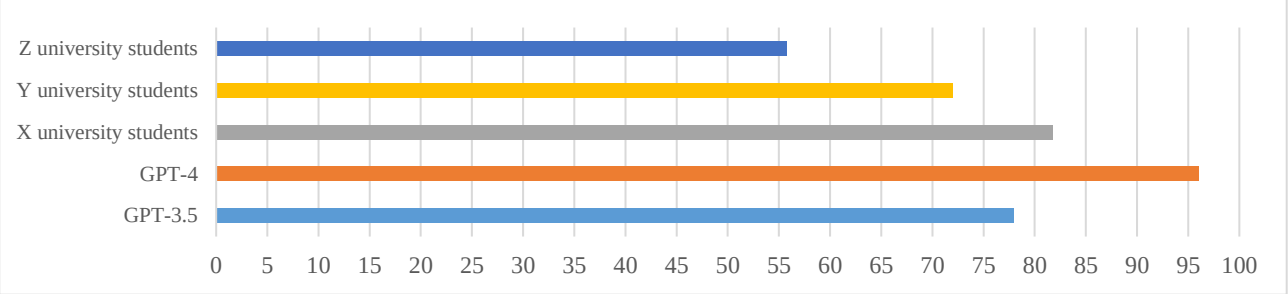

Figure 3. Total scores of the MCQ exams

Source: Elaborated by Authors

Firstly, the performance of ChatGPT-4 is outstanding. This underlines the advanced capabilities of this latest AI model to accurately process and answer a wide variety of questions. ChatGPT-4's proficiency in these exams demonstrates that AI technology has reached a level of sophistication where it can competently handle diverse and complex information, a crucial element in educational environments. The overall performance of students from X

university is superior to that of ChatGPT-3.5. While the artificial intelligence demonstrates competence in answering test questions, students from X University exhibit a deeper understanding and greater cognitive flexibility than ChatGPT-3.5. The data presented in Figure 2. provides insights into the diversity of human learning, as evidenced by the variations in total scores among students from different universities.

The exam of open-ended questions

Questions 1 and 2: After a short scenario of 120 words about reasons for the preference of tourist guides and their characteristics, two questions were asked to ChatGPT and students. The first question is: *"What do agencies look for when selecting tourist guides?"* and the second question is: *"What are the essential skills and qualities that a tourist guide must have?"*

ChatGPT-3.5's answer to the first question was short and clear. It stated that travel agencies look for the following characteristics when selecting tourist guides: "knowledge, language skills, interpersonal skills, cultural sensitivity, professionalism, safety awareness, storytelling, and flexibility." It listed these characteristics in bullet points and provided quite brief explanations for each. ChatGPT-4 has identified several essential qualities and qualifications that travel agencies look for, such as "knowledge, effective communication, strong interpersonal abilities, competent problem-solving skills, genuine enthusiasm and passion, adaptability and flexibility, professionalism, relevant certifications, reputation, references, and cultural sensitivity." ChatGPT-4's answer and explanations were quite satisfactory.

Similarly, ChatGPT-3.5 gave a short and clear answer to the second question. Its explanations were quite short. According to ChatGPT-3.5, the characteristics that tourist guides should have been as follows: "extensive knowledge, language proficiency, communication skills, interpersonal skills, flexibility, problem-solving, cultural sensitivity, safety awareness, professionalism, and enthusiasm." In the answer to the second question, ChatGPT-4, unlike ChatGPT-3.5, added the following: "Charismatic, patient, stamina, and attention to detail." ChatGPT-4's answer and explanations were more detailed.

When we examined the answers of university students to the first and second questions, we observed that all university students provided deeper and more detailed answers than ChatGPT-3.5. Particularly, the persuasive and detailed responses of students from X University were noteworthy. Also, the answers of some students from other universities were better than ChatGPT-4. The fact that some students from all universities had previously worked as assistant staff on tours accompanied by tourist guides could have influenced this outcome. Practical training improves academic knowledge and enables learners to provide answers that are both theoretically sound and based on real-world experience. The students' better answers appear to be due to this blend of academic learning and practical experience.

Questions 3 and 4: We gave a scenario to ChatGPT and students. This scenario explains the perspective of an experienced tour guide about the various preferences of tourists, including people who get pleasure from socializing and leisure activities, as well as those who have a strong inclination towards native customs and traditions. The scenario includes problems such as effectively managing expectations around shopping and additional fees, opinions, and beliefs. In addition, the different reactions of tourists with different religious beliefs to the guide's neutral narratives on the history of religions, etc. are also included in the scenario. First questions is: "Tourists who are very interested in customs and traditions can be included in which tourist typologies?" and the second question is: "Can tourists make incorrect inferences from the guide's narratives? What can be done to prevent this?"

The answer given by ChatGPT 3.5 to the first question is both clear and contextually appropriate. Describing these tourists as 'culture tourists' is consistent with the findings of tourism studies, but answer was too short and absent of explanation. ChatGPT-4 gave a more detailed answer. It listed 5 typologies: "Cultural tourist, heritage tourist, ethno-tourist, educational tourist, and experiential tourist." It also gave a detailed explanation for each of these typologies. Most of the university students wrote "cultural tourist" as their answer. Also, some of them added the other typologies. Generally, the answers of university students were more descriptive than ChatGPT3.5. Many of them were able to write answers as deep and clear as ChatGPT-4.

The answer of ChatGPT 3.5 to the second question was comprehensive and perceptive. The solution-focused approach, which advocates for the use of clear and precise language, providing relevant context, promoting the asking of questions, maintaining neutrality, showing respect for varied perspectives, utilizing visual aids, and monitoring feedback, is acceptable. The approach of ChatGPT-4 to the problem of preventing incorrect inferences from guide narratives is comprehensive and practical. The answer recognizes multiple factors that can lead to misunderstandings, including differences in culture and language obstacles. It suggests effective approaches such as using clear communication, providing contextual information, being aware of the audience, keeping neutrality, encouraging questions, implementing feedback mechanisms, providing cultural sensitivity training, and participating in ongoing professional growth. The suggestions show a deep understanding of the key elements of effective guidance and audience engagement.

Still, most of the answers from the students were better than ChatGPT 3.5. Not all of the students could give as thorough answers as they did in ChatGPT 4, though. More detailed and useful answers came from students who had

experience as a tourist guide. Students with more experience gave specific examples and points of view that were based on real-life scenarios by working directly with travelers and gaining hands-on experience. A new point of view might come from having experience in the field. Because of the human information gained in real life situations, the answers were completer and more complicated than those given by AI.

5. Discussion

The study's results indicate that ChatGPT-4's proficiency in multiple-choice questions (MCQs) is consistent with AI's capacity to handle structured, knowledge-based inquiries. This is consistent with the results of Dwivedi et al. (2023) in their exhaustive analysis of AI tools, including ChatGPT, from a variety of disciplinary perspectives. It was noted that these AI programs are effective in improving efficiency and performing well on duties that necessitate access to extensive databases. This capability can be ascribed to the AI's capacity for systematic retrieval of information and data processing. ChatGPT possesses attributes that render it an intriguing instrument for assisting in the theoretical aspects of instructing tourist guides in regions that necessitate descriptive knowledge.

The students, those with experience in tourist guiding, demonstrated higher scores on the open-ended questions than the ChatGPT-4. The human element in education is underscored by this outcome, which is in accordance with a systematic review conducted by Crompton & Burke (2023). The research cautions that, although AI can aid in educational objectives, it should not be capable of delivering the high-level knowledge that only humans can provide in areas that necessitate hands-on learning and analytical skills. The research indicates that the empirical wisdom acquired through tourist guide programs is not comprehended by sophisticated models like ChatGPT-4. Furthermore, Alawida et al. (2023) emphasizes the challenges of ChatGPT in tasks that are typically acquired through such experience and involve hands-on practical knowledge. ChatGPT is one of the artificial intelligence systems that can be employed to analyze and produce information from data. It is inadequate to apply this knowledge in a practical setting. This deficiency emphasizes the importance of practical experience.

The findings are indeed analogous with those obtained from other earlier studies on AI, which, even while the epic information dissemination as well as learning with the aid of AI, is not fully capable of replacing or substituting human beings' experience and knowledge. The findings are parallel to those in other works by Ghosh et al., 2023; Michalon & Camacho-Zuñiga, 2023; Owan et al., 2023; Vasconcelos & Dos Santos, 2023. This coherence would only reiterate how artificial intelligence, specifically in second-order versions such as ChatGPT-4, is best used as a tool for support rather than a sole medium for instruction in pedagogy. Artificial intelligence can be used in the course of the tourist guide or in which the theoretical information is handed over properly. This is beneficial because it enables one to deal effectively with such extensive and diverse knowledge as found in the profession of tourist guides. suggesting that although AI may greatly assist in the transmission of information and learning, it is not capable of completely mimicking human experience and comprehension. This consistency supports the idea that AI, especially sophisticated models such as ChatGPT-4, is better suited as a supplementary tool rather than an alternative in educational environments.

Integrating of AI helps to enhance the student experience in tourist guiding because the system will become supportive with the delivery of tailored information, including MCQs and interactive modules. AI solutions pay attention to the needs of students and their preferred learning methods. For instance, AI can deliver personalized information through multiple-choice questions and interactive modules, enhancing the overall learning experience for students studying tourist guiding. This approach is crucial for effectively managing the extensive and varied knowledge required in this field.

Artificial intelligence lacks the deep understanding and ability to teach that human instructors have. Especially in a field such as tourist guiding, where cultural and ethical issues and personal experiences are intense, it does not seem possible to provide training based entirely on artificial intelligence.

Artificial intelligence (AI) is highly efficient in handling jobs that require a large amount of data since it can provide useful feedback and support customized learning experiences. Conversely, AI may not be able to fully satisfy the cognitive abilities, innovative thinking, and emotional understanding that human trainers possess. Proficiency in these abilities is especially crucial in the tourism industry, since it heavily relies on interpersonal communication and cross-cultural comprehension. Hence, a collaborative strategy is necessary, whereby AI technologies and human-centric teaching techniques coexist and enhance each other.

6. Conclusion

ChatGPT-4 outperformed both ChatGPT-3.5 and university students in the multiple-choice examination in the field of tourist guiding, getting a score of 96 out of 100. Despite ChatGPT-3.5 falling behind ChatGPT-4, it showed acceptable results, displaying a moderate degree of achievement. University students, namely those from University X, exhibited a diverse range of results; they outperformed ChatGPT-3.5 in terms of overall scoring but fell short of the performance level achieved by ChatGPT-4.

ChatGPT-4 outperformed ChatGPT-3.5 in open-ended question types by delivering deeper and comprehensive responses, while ChatGPT-3.5 tended to produce brief and short answers. When analyzing the comments of university students, it was noted that those with practical expertise in

the sector offered more profound and persuasive answers compared to ChatGPT-4. This scenario highlights the inherent significance of human expertise and hands-on know-how in the domain, which cannot be completely duplicated by artificial intelligence.

In summary, the study demonstrates that AI, namely ChatGPT-4, performs exceptionally well in MCQs that rely on knowledge, while human learners exhibit a superior level of comprehension and practical application in open-ended scenarios. These findings indicate that AI could be used in educational settings in the tourism area. AI can be a useful tool for sharing knowledge and evaluating learning, in addition to traditional human-centered teaching methods.

Theoretical Implications

The most significant contribution of this paper lies in the detailed review of the strengths and weaknesses of AI in the sphere of tourist guidance study, which enhances the theoretical framework. This kind of evaluation is helpful for academicians and educators who are considering making use of artificial intelligence in their training programs for tourist guides. The results of this study add to the literature on AI in tourism, inferring when human intervention is required within the context of scenario-based application areas.

This study demonstrates the potential of advanced models like ChatGPT in disseminating both theoretical and factual information. It lays the groundwork for future research on incorporating artificial intelligence into tourist guiding and tourism education environments. This approach helps understand how AI can be integrated into traditional educational methods. The introduction of AI has the potential to transform tourism education and research by enhancing and expanding learning experiences. According to Ouyang & Jiao (2021), AI can serve as a supplementary tool in tourism education, working alongside learners to support their unique learning processes. This perspective is based on a learning model that combines cognitive and social constructivism.

This study demonstrates that while AI has developed to a competent level in processing theoretical knowledge, the human educator remains crucial in teaching practical comprehension and application abilities. Suggesting a collaborative strategy in which AI and human educators serve in parallel roles provides evidence for the theoretical implications for tourist guide education. More concretely, it can contribute to a better learning process or theoretical education in AI with practical experiences provided by the human educator.

The study assists in affecting educational development theories for tourism approaches in the digital age. This possibly presents the beginning of new pedagogical concepts for a flexible and adaptive way to education in tourism and guidance of tourists, based on the use of newly introduced methods, such as artificial intelligence.

Practical Implications

ChatGPT can generate ideal knowledge-based responses, making it promising for creating training material and digital applications for tourist guides. AI capabilities facilitate self-learning, allowing guides to research and relate to their profession. These tools effectively assess information on critical topics such as cultural heritage, local history, and geography.

The study concludes that tourism education should include a mix of AI-driven tools and human instructors. AI tools will handle theoretical learning, while human instructors focus on sharing practical experiences. This approach will provide more engaging and comprehensive learning for students. The study's conclusions will influence the development of educational courses for the tourism sector. AI informs educators about creating more effective and well-rounded courses, recognizing that practical skills require human involvement.

Offering personalized information, ChatGPT can help develop learning experiences for students, addressing their diverse learning needs and cultural backgrounds in tourism education. Knowledge embedded in AI technologies within education will prepare future tourist guides to use technology effectively in their professions.

Further Studies

This study indicates that further studies are needed to understand the impact of ChatGPT on the learning of tourism. Further studies should demonstrate that course assessment through the intervention of ChatGPT would be viable (Skavronskaya et al., 2023). There is a need for studies and development in implementing the examination preparation and assessment using AI with effectiveness (Gimpel et al., 2023). Though tourism study and education are still developing and diversifying, at the moment, with the future being reformed, AI seems to be an important future for it (Ivanov & Soliman, 2023).

This study would prove the performance of ChatGPT in training tourist guides, which can be a valuable base reference for future researchers. What follows is the identification of what AI can do in terms of its strengths and, at the same time, its limitations. The current study paves the way for further investigations into this framework approach to tourism education. We consider it a significant milestone for future research and a roadmap for the potential advancement of AI technologies in the tourism literature. This is the same study that lends itself to the theories of AI fields demanding empathy, ethical consideration, and cultural understanding—all of which form the core of tourism studies. This paves the way for further research into the boundaries of AI in areas that are very dependent on human potentialities and actualizations.

In the future, large-scale studies involving people from a variety of organizations are needed, whose results are applicable to a wider range of situations. This could help

us learn more about how AI tools like ChatGPT are used in different cultural and educational settings. On the other hand, studying the effectiveness of other AI models or new versions of ChatGPT will yield valuable insights into betterment and construction in the area of tourism education. It is needed to be a rich source of solid knowledge on the differences in effectiveness and application of different AI techniques. Lastly, all future studies should be able to see through the ethical issues surrounding the use of AI in tourism education—such as data privacy, academic dishonesty, and the potential ripple effect on certain roles previously attributed to teachers.

References

Abduljabbar, R., Dia, H., Liyanage, S., & Bagloee, S. A. (2019). Applications of artificial intelligence in transport: an overview. *Sustainability*, *11*(1), 189. https://doi.org/10.3390/su11010189

Agarwal, S., Agarwal, B., & Gupta, R. (2022). Chatbots and virtual assistants: A bibliometric analysis. *Library Hi Tech*, *40*(4), 1013–1030. https://doi.org/10.1108/LHT-09-2021-0330

Alawida, M., Mejri, S., Mehmood, A., Chikhaoui, B., & Isaac Abiodun, O. (2023). A comprehensive study of chatgpt: Advancements, limitations, and ethical considerations in natural language processing and cybersecurity. *Information*, *14*(8), 462. https://doi.org/10.3390/info14080462

Alazaizeh, M. M., Jamaliah, M. M., Mgonja, J. T., & Ababneh, A. (2019). Tour guide performance and sustainable visitor behavior at cultural heritage sites. *Journal of Sustainable Tourism*, *27*(11), 1708–1724. https://doi.org/10.1080/09669582.2019.1658766

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

Athavale, A., Baier, J., Ross, E., & Fukaya, E. (2023). The potential of chatbots in chronic venous disease patient management. *JVS-Vascular Insights*, 100019. https://doi.org/10.1016/j.jvsvi.2023.100019

Atlas, S. (2023). ChatGPT for higher education and professional development: A guide to conversational AI. *College of Business Faculty Publications*. https://digitalcommons.uri.edu/cba_facpubs/548

Aydın, Ö., & Karaarslan, E. (2023). Is ChatGPT leading generative AI? What is beyond expectations? *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4341500

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

Bommarito II, M., & Katz, D. M. (2022). GPT Takes the Bar Exam. *arXiv*. http://arxiv.org/abs/2212.14402

Cao, Y., & Zhai, J. (2023). Bridging the gap – the impact of ChatGPT on financial research. *Journal of Chinese Economic and Business Studies*, *21*(2), 177–191. https://doi.org/10.1080/14765284.2023.2212434

Chang, C.-Y., Kuo, S.-Y., & Hwang, G.-H. (2022). Chatbot-facilitated nursing education: Incorporating a knowledge-based chatbot system into a nursing training program. *Educational Technology & Society*, 25(1), 15-27.

Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, A. (2018). Artificial Intelligence trends in education: A narrative overview. *Procedia Computer Science*, *136*, 16–24. https://doi.org/10.1016/j.procs.2018.08.233

Chen, L., Chen, P., & Lin, Z. (2020). Artificial Intelligence in education: A Review. *IEEE Access*, *8*, 75264–75278. https://doi.org/10.1109/ACCESS.2020.2988510

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

Chui, K. T., Lytras, M. D., & Visvizi, A. (2018). Energy sustainability in smart cities: Artificial intelligence, smart monitoring, and optimization of energy consumption. *Energies*, *11*(11), 2869. https://doi.org/10.3390/en11112869

Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education*, *20*(1), 22. https://doi.org/10.1186/s41239-023-00392-8

Currie, G., Singh, C., Nelson, T., Nabasenja, C., Al-Hayek, Y., & Spuur, K. (2023). ChatGPT in medical imaging higher education. *Radiography*, *29*(4), 792–799. https://doi.org/10.1016/j.radi.2023.05.011

Dalalah, D., & Dalalah, O. M. A. (2023). The false positives and false negatives of generative AI detection tools in education and academic research: The case of ChatGPT. *The International Journal of Management Education*, *21*(2), 100822. https://doi.org/10.1016/j.ijme.2023.100822

Dias, J., Kamdi, D., Gharat, N., & Chudhari, P. (2019). Chatbot for government examination using AI. *Journal of Engineering*, *9*(4).

Dinesh, A., Kamal, B., Akash, M., & Surthik, K. R. (2023). Implementation of artificial intelligence to the prediction of the mechanical properties of concrete: A review. *Materials Today: Proceedings*, S221478532302727X. https://doi.org/10.1016/j.matpr.2023.05.101

Doborjeh, Z., Hemmington, N., Doborjeh, M., & Kasabov, N. (2022). Artificial intelligence: A systematic review of methods and applications in hospitality and tourism. *International Journal of Contemporary Hospitality Management*, *34*(3), 1154–1176. https://doi.org/10.1108/IJCHM-06-2021-0767

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., … Wright, R. (2023). Opinion Paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational

AI for research, practice and policy. *International Journal of Information Management*, *71*, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

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

Fijačko, N., Gosak, L., Štiglic, G., Picard, C. T., & John Douma, M. (2023). Can ChatGPT pass the life support exams without entering the American heart association course? *Resuscitation*, *185*, 109732. https://doi.org/10.1016/j.resuscitation.2023.109732

Geerling, W., Mateer, G. D., Wooten, J., & Damodaran, N. (2023). ChatGPT has mastered the principles of economics: Now What? *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4356034

Ghosh, A., Maini Jindal, N., Gupta, V. K., Bansal, E., Kaur Bajwa, N., & Sett, A. (2023). Is ChatGPT's knowledge and ınterpretatıve ability comparable to first professional mbbs (bachelor of medicine, bachelor of surgery) students of ındia in taking a medical biochemistry examination? *Cureus*. https://doi.org/10.7759/cureus.47329

Giannos, P., & Delardas, O. (2023). Performance of ChatGPT on UK standardized admission tests: Insights from the BMAT, TMUA, LNAT, and TSA Examinations. *JMIR Medical Education*, *9*, e47737. https://doi.org/10.2196/47737

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 ımplications of large language models for medical education and knowledge assessment [Preprint]. Medical Education. https://doi.org/10.1101/2022.12.23.22283901

Gimpel, H., Hall, K., Decker, S., Eymann, T., Lämmermann, L., Mädche, A., Röglinger, M., Ruiner, C., Schoch, M., Schoop, M., Urbach, N., & Vandirk, S. (2023). Unlocking the power of generative aı models and systems such as gpt-4 and chatgpt for higher education. *Hohenheim Discussion Papers in Business, Economics and Social Sciences*. https://doi.org/10.13140/RG.2.2.20710.09287/2

Hartmann, J., Schwenzow, J., & Witte, M. (2023). The political ideology of conversational AI: Converging evidence on ChatGPT's pro-environmental, left-libertarian orientation. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4316084

Ho, W. L. J., Koussayer, B., & Sujka, J. (2023). ChatGPT: Friend or foe in medical writing? An example of how ChatGPT can be utilized in writing case reports. *Surgery in Practice and Science*, *14*, 100185. https://doi.org/10.1016/j.sipas.2023.100185

Iqbal, M. Z., & Faiz, M. F. I. (2020). Active surveillance for Covıd-19 through artificial intelligence using real-time speech-recognition mobile application. *2020 IEEE International Conference on Consumer Electronics - Taiwan (ICCE-Taiwan)*, 1–2. https://doi.org/10.1109/ICCE-Taiwan49838.2020.9258276

Ivanov, S., & Soliman, M. (2023). Game of algorithms: ChatGPT implications for the future of tourism education and research. *Journal of Tourism Futures*, *9*(2), 214–221. https://doi.org/10.1108/JTF-02-2023-0038

Jenneboer, L., Herrando, C., & Constantinides, E. (2022). The impact of chatbots on customer loyalty: a systematic literature review. *Journal of Theoretical and Applied Electronic Commerce Research*, *17*(1), 212–229. https://doi.org/10.3390/jtaer17010011

Jin, J. Q., & Dobry, A. S. (2023). ChatGPT for healthcare providers and patients: Practical implications within dermatology. *Journal of the American Academy of Dermatology*, S0190962223011064. https://doi.org/10.1016/j.jaad.2023.05.081

Keiper, M. C., Fried, G., Lupinek, J., & Nordstrom, H. (2023). Artificial intelligence in sport management education: Playing the AI game with ChatGPT. *Journal of Hospitality, Leisure, Sport & Tourism Education*, *33*, 100456. https://doi.org/10.1016/j.jhlste.2023.100456

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, 003368822311628. https://doi.org/10.1177/00336882231162868

Kooli, C. (2023). Chatbots in education and research: A critical examination of ethical implications and solutions. *Sustainability*, *15*(7), 5614. https://doi.org/10.3390/su15075614

Kortemeyer, G. (2023). Could an Artificial-Intelligence agent pass an introductory physics course? *Physical Review Physics Education Research*, *19*(1), 010132. https://doi.org/10.1103/PhysRevPhysEducRes.19.010132

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

Kutyauripo, I., Rushambwa, M., & Chiwazi, L. (2023). Artificial intelligence applications in the agrifood sectors. *Journal of Agriculture and Food Research*, *11*, 100502. https://doi.org/10.1016/j.jafr.2023.100502

Lee, D., & Yoon, S. N. (2021). Application of artificial ıntelligence-based technologies in the healthcare industry: Opportunities and Challenges. *International Journal of Environmental Research and Public Health*, *18*(1), 271. https://doi.org/10.3390/ijerph18010271

Leiter, C., Zhang, R., Chen, Y., Belouadi, J., Larionov, D., Fresen, V., & Eger, S. (2023). ChatGPT: A Meta-Analysis after 2.5 Months. *arXiv*. http://arxiv.org/abs/2302.13795

Leung, X. Y., & Wen, H. (2020). Chatbot usage in restaurant takeout orders: A comparison study of three ordering methods. *Journal of Hospitality and Tourism*

Management, *45*, 377–386. https://doi.org/10.1016/j.jhtm.2020.09.004

Li, L. (2023). Role of chatbots on gastroenterology: Let's chat about the future. *Gastroenterology & Endoscopy*, S2949752323000341. https://doi.org/10.1016/j.gande.2023.06.002

Li, M., & Wang, R. (2023). Chatbots in e-commerce: The effect of chatbot language style on customers' continuance usage intention and attitude toward brand. *Journal of Retailing and Consumer Services*, *71*, 103209. https://doi.org/10.1016/j.jretconser.2022.103209

Li, W., Wang, J., Yan, Y., Yuan, P., Cao, C., Li, S., & Wu, Q. (2023). Potential applications of ChatGPT in endoscopy: Opportunities and limitations. *Gastroenterology & Endoscopy*, S294975232300033X. https://doi.org/10.1016/j.gande.2023.06.001

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

Lin, X., Shao, B., & Wang, X. (2022). Employees' perceptions of chatbots in B2B marketing: Affordances vs. disaffordances. *Industrial Marketing Management*, *101*, 45–56. https://doi.org/10.1016/j.indmarman.2021.11.016

Lin, Y., & Yu, Z. (2023). A bibliometric analysis of artificial intelligence chatbots in educational contexts. *Interactive Technology and Smart Education*. https://doi.org/10.1108/ITSE-12-2022-0165

Luo, B., Lau, R. Y. K., & Li, C. (2023). Emotion-regulatory chatbots for enhancing consumer servicing: An interpersonal emotion management approach. *Information & Management*, *60*(5), 103794. https://doi.org/10.1016/j.im.2023.103794

Ma, Y. (2023). The potential application of ChatGPT in gastrointestinal pathology. *Gastroenterology & Endoscopy*, *1*(3), 130–131. https://doi.org/10.1016/j.gande.2023.05.002

Melián-González, S., Gutiérrez-Taño, D., & Bulchand-Gidumal, J. (2021). Predicting the intentions to use chatbots for travel and tourism. *Current Issues in Tourism*, *24*(2), 192–210. https://doi.org/10.1080/13683500.2019.1706457

Michalon, B., & Camacho-Zuñiga, C. (2023). ChatGPT, a brand-new tool to strengthen timeless competencies. *Frontiers in Education*, *8*, 1251163. https://doi.org/10.3389/feduc.2023.1251163

Naser, M. Z., Ross, B., Ogle, J., Kodur, V., Hawileh, R., Abdalla, J., & Thai, H.-T. (2023). Can AI Chatbots Pass the Fundamentals of Engineering (FE) and Principles and Practice of Engineering (PE) Structural Exams? *arXiv*. https://doi.org/10.48550/ARXIV.2303.18149

Newton, P. M., & Xiromeriti, M. (2023). Chatgpt performance on mcq exams in higher education. a pragmatic scoping review [Preprint]. EdArXiv. https://doi.org/10.35542/osf.io/sytu3

Nunes, D., Primi, R., Pires, R., Lotufo, R., & Nogueira, R. (2023). Evaluating Gpt-3.5 and Gpt-4 models on brazilian university admission exams (arXiv:2303.17003). arXiv. http://arxiv.org/abs/2303.17003

Okonkwo, C. W., & Ade-Ibijola, A. (2021). Chatbots applications in education: A systematic review. *Computers and Education: Artificial Intelligence*, *2*, 100033. https://doi.org/10.1016/j.caeai.2021.100033

OpenAI. (n.d.). *GPT-4*. OpenAI Research. https://openai.com/research/gpt-4

OpenAIb. (n.d.). *About OpenAI*. https://openai.com/about

Orden-Mejía, M. A., & Huertas, A. (2022). Tourist interaction and satisfaction with the chatbot evokes pre-visit destination image formation? A case study. *Anatolia*, 1–15. https://doi.org/10.1080/13032917.2022.2072918

Ouyang, F., & Jiao, P. (2021). Artificial intelligence in education: The three paradigms. *Computers and Education: Artificial Intelligence*, *2*, 100020. https://doi.org/10.1016/j.caeai.2021.100020

Owan, V. J., Abang, K. B., Idika, D. O., Etta, E. O., & Bassey, B. A. (2023). Exploring the potential of artificial intelligence tools in educational measurement and assessment. *Eurasia Journal of Mathematics, Science and Technology Education*, *19*(8), em2307. https://doi.org/10.29333/ejmste/13428

Parsons, H., Houge Mackenzie, S., & Filep, S. (2019). Facilitating self-development: How tour guides broker spiritual tourist experiences. *Tourism Recreation Research*, *44*(2), 141–152. https://doi.org/10.1080/02508281.2019.1582159

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

Peters, M. A., Jackson, L., Papastephanou, M., Jandrić, P., Lazaroiu, G., Evers, C. W., Cope, B., Kalantzis, M., Araya, D., Tesar, M., Mika, C., Chen, L., Wang, C., Sturm, S., Rider, S., & Fuller, S. (2023). AI and the future of humanity: ChatGPT-4, philosophy and education – Critical responses. *Educational Philosophy and Theory*, 1–35. https://doi.org/10.1080/00131857.2023.2213437

Pikhart, M. (2020). Intelligent information processing for language education: The use of artificial intelligence in language learning apps. *Procedia Computer Science*, *176*, 1412–1419. https://doi.org/10.1016/j.procs.2020.09.151

Pursnani, V., Sermet, Y., & Demir, I. (2023). Performance of ChatGPT on the US Fundamentals of Engineering Exam: Comprehensive Assessment of Proficiency and Potential Implications for Professional Environmental Engineering Practice. *ArXiv*. https://doi.org/10.48550/ARXIV.2304.12198

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning & Teaching*, *6*(1). https://doi.org/10.37074/jalt.2023.6.1.9

Short, C. E., & Short, J. C. (2023). The artificially intelligent entrepreneur: ChatGPT, prompt engineering, and entrepreneurial rhetoric creation. *Journal of Business Venturing Insights*, *19*, e00388. https://doi.org/10.1016/j.jbvi.2023.e00388

Singh, S. V., & Hiran, K. K. (2022). The Impact of AI on Teaching and learning in higher education technology. *Journal of Higher Education Theory and Practice*, *22*(13). https://doi.org/10.33423/jhetp.v22i13.5514

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

Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing*, *57*, 100752. https://doi.org/10.1016/j.asw.2023.100752

Susnjak, T. (2022). *ChatGPT: The End of Online Exam Integrity?* (arXiv:2212.09292). arXiv. http://arxiv.org/abs/2212.09292

Tsaur, S.-H., & Teng, H.-Y. (2017). Exploring tour guiding styles: The perspective of tour leader roles. *Tourism Management*, *59*, 438–448. https://doi.org/10.1016/j.tourman.2016.09.005

TUREB. (n.d.-a). *Şanlıurfa Regional Certificate Program General Culture Exam Questions*. https://www.tureb.org.tr/tr/Page/Detail/93

TUREB. (n.d.-b). *Trabzon Regional Certificate Program General Culture Exam Questions*. https://www.tureb.org.tr/tr/Page/Detail/92

Tyson, M. M., & Sauers, N. J. (2021). School leaders' adoption and implementation of artificial intelligence. *Journal of Educational Administration*, *59*(3), 271–285. https://doi.org/10.1108/JEA-10-2020-0221

Vasconcelos, M. A. R., & Dos Santos, R. P. (2023). Enhancing STEM learning with ChatGPT and Bing Chat as objects to think with: A case study. *Eurasia Journal of Mathematics, Science and Technology Education*, *19*(7), em2296. https://doi.org/10.29333/ejmste/13313

Victor, B. G., Kubiak, S., Angell, B., & Perron, B. E. (2023). Time to move beyond the aswb licensing exams: Can generative artificial ıntelligence offer a way forward for social work? *Research on Social Work Practice*, *33*(5), 511–517. https://doi.org/10.1177/10497315231166125

Wang, C., Li, Y., Fu, W., & Jin, J. (2023). Whether to trust chatbots: Applying the event-related approach to understand consumers' emotional experiences in interactions with chatbots in e-commerce. *Journal of Retailing and Consumer Services*, *73*, 103325. https://doi.org/10.1016/j.jretconser.2023.103325

Wang, F., Wong, W.-K., Reivan Ortiz, G. G., Shraah, A. A., Mabrouk, F., Li, J., & Li, Z. (2023). Economic analysis of sustainable exports value addition through natural resource management and artificial intelligence. *Resources Policy*, *82*, 103541. https://doi.org/10.1016/j.resourpol.2023.103541

Xu, Y., Niu, N., & Zhao, Z. (2023). Dissecting the mixed effects of human-customer service chatbot interaction on customer satisfaction: An explanation from temporal and conversational cues. *Journal of Retailing and Consumer Services*, *74*, 103417. https://doi.org/10.1016/j.jretconser.2023.103417

Zaitsu, W., & Jin, M. (2023). *Distinguishing ChatGPT(-3.5, -4)-generated and human-written papers through Japanese stylometric analysis*.

Zhang, C., & Lu, Y. (2021). Study on artificial intelligence: The state of the art and future prospects. *Journal of Industrial Information Integration*, *23*, 100224. https://doi.org/10.1016/j.jii.2021.100224

Zhu, D., & Xu, H. (2021). Guides' handling of multiple roles in interactions with tourists: The prism of impression management. *Tourism Management*, *85*, 104306. https://doi.org/10.1016/j.tourman.2021.104306

Jomat

Journal of multidisciplinary academic tourism

2024, 9 (3): 231-243

https://doi.org/10.31822/jomat.2024-9-3-231

ISSN: 2645-9078

INFO PAGE

The performance of artificial intelligence in the exams of tourist guidance

Abstract

The aim of this study is to evaluate the efficiency of ChatGPT versions 3.5 and 4 for training tourist guides. The study followed a systematic approach by conducting assessments on undergraduate students from three institutions who are enrolled in tourist guide education programs and both ChatGPT versions. Competent academicians assessed a succession of questions in the form of open-ended and multiple-choice questions. The mean scores obtained on the multiple-choice test for ChatGPT-4 were better than those of both ChatGPT-3.5 and college students, thereby indicating that ChatGPT-4 has greatly improved. Nevertheless, when responding to open-ended queries, individuals with real-life experience as tour guides gave much more inclusive as well as convincing answers compared to ChatGPT-4. This underscores the importance of hands-on experiences in training tour guides, where AI technology is currently weak. This study contributes to better comprehension regarding the role played by artificial intelligence (AI) in education with reference to the tourism industry specifically. While at the same time emphasizing how critical human expertise is needed during practical learning sessions, this implies that AI has potential for disseminating theoretical knowledge. The results suggest that AI is a beneficial supplementary aid in educational environments, rather than a replacement for human-centered instructional approaches.

Keywords: *Artificial intelligence, Large language models (LLMs), Tourism education, Tourist guidance, Exam.*

Authors

Full Name	Author contribution roles	Contribution rate
Abdullah Ülkü:	*Conceptualism, Methodology, Validation, Formal Analysis, Investigation, Resources, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization*	*100%*

Author statement: *Author(s) declare(s) that All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.* ***Declaration of Conflicting Interests:*** *The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article*

Ethics Committee Satatement: Ethics committee report is available for this research and it has been documented to the journal

Ethics committee: **Harran University Social Sciences and Humanities Ethics Committee**

Date of ethics committee decision: **45218**

Ethics committee decision number: **2023/150**