İZU Journal of Education / e-ISSN: 2667-7938
(June 2026), Volume 8, No 1, 1–26
https://doi.org/10.46423/izujed.1793325

ZAIMPRESS

RESEARCH

Pre-Service Teachers' Experiences with Using Chatgpt for Mathematical Tasks

Zeynep Çiftçi [1] Tuba Öz [1] Orhan Çiftçi [2]

[1] Asst. Prof. Dr., Atatürk University, Department of Mathematics and Science Education, Erzurum, Türkiye
[2] Ph.D., Republic of Türkiye Ministry of National Education, Erzurum, Türkiye

Received: 02 October 2025 / Accepted: 06 February 2026 / Published online: 30 June 2026

Abstract
This study focuses on determining how pre-service teachers create prompts when using ChatGPT to solve mathematics tasks, the accuracy of the answers of pre-service teachers and ChatGPT, and the opinions of pre-service teachers about ChatGPT. The potential of ChatGPT to be used in education is interesting because large language models such as ChatGPT inherently provide different answers to each user in line with the prompts entered and these answers carry the risk of being inaccurate. This study is thought to contribute to the expanding literature on whether ChatGPT can be used in the field of mathematics education. 32 pre-service teachers were asked to complete two distinct mathematical tasks with the support of ChatGPT. ChatGPT chats, task solutions, and opinion forms of pre-service teachers were analyzed using qualitative content analysis. According to the results of the study, while the prompts created by the pre-service teachers in ChatGPT showed a great variety, most of the pre-service teachers created prompts in ChatGPT by writing the entire task. It was concluded that although the ChatGPT answers were correct at an acceptable level in algebra, they were incorrect at a very high rate in geometry. Among the opinions of pre-service teachers about ChatGPT, quick access to information and feedback, sufficient explanatory information, lack of figures and visuals, inconsistent information, and being deficient in the field of mathematics come to the fore.

Keywords: ChatGPT, ChatBot, OpenAI, AIEd, mathematics education, artificial intelligence

To cite: Çiftçi, Z., & Öz, T., & Çiftçi O. (2026). Pre-Service Teachers' Experiences with Using Chatgpt for Mathematical Tasks. *İZÜ Eğitim Dergisi, 8*(1), 1–26. https://doi.org/10.46423/izujed.1793325

Corresponding author: Tuba Öz tkaplan@atauni.edu.tr

This article is open access and distributed under the terms of the CC BY-NC 4.0 license.

İZÜ Eğitim Dergisi / e-ISSN: 2667-7938
(Haziran 2026), Cilt 8, Sayı 1, 1–26
https://doi.org/10.46423/izujed.1793325

ZAIMPRESS

ARAŞTIRMA

Öğretmen Adaylarının ChatGPT'yi Matematiksel Görevlerde Kullanımına İlişkin Deneyimleri

Zeynep Çiftçi [1] **Tuba Öz** [1] **Orhan Çiftçi** [2]

[1] Dr.Öğrt.Üyesi, Atatürk Üniversitesi, Matematik ve Fen Bilimleri Eğitimi Bölümü, Erzurum, Türkiye
[2] Dr., TC Millî Eğitim Bakanlığı, Erzurum, Türkiye

Başvuru: 02 Ekim 2025 / Kabul: 06 Şubat 2026 / Online Yayın: 30 Haziran 2026

Öz

Bu çalışma, öğretmen adaylarının matematik görevlerini çözümünde ChatGPT kullanırken nasıl istem oluşturduklarını, öğretmen adaylarının ve ChatGPT'nin cevaplarının doğruluğunu ve öğretmen adaylarının ChatGPT ile ilgili görüşlerini belirlemeye odaklanmaktadır. ChatGPT gibi büyük dil modelleri doğası gereği her kullanıcıya girilen istemler doğrultusunda farklı cevaplar sunması ve bu cevapların hatalı olma riski taşıması nedeniyle ChatGPT'nin eğitimde kullanılma potansiyeli ilgi çekicidir. Bu çalışma ile ChatGPT'nin matematik eğitimi alanında kullanılıp kullanılamayacağı konusunda genişleyen literatüre katkı sağlayacağı düşünülmektedir. 32 öğretmen adayından ChatGPT yardımıyla iki farklı matematik görevini çözmeleri istenmiştir. Öğretmen adaylarının ChatGPT sohbetleri, görev çözümleri ve görüş formları nitel içerik analizi kullanılarak analiz edilmiştir. Araştırmanın sonuçlarına göre öğretmen adaylarının ChatGPT'de oluşturdukları istemler büyük bir çeşitlilik gösterirken öğretmen adaylarının çoğu görevin tamamını yazarak ChatGPTde istem oluşturmuşlardır. ChatGPT cevaplarının cebir alanında kabul edilebilir düzeyde doğru olmasına karşın geometri alanında oldukça yüksek oranda yanlış olduğu sonucuna ulaşılmıştır. Öğretmen adaylarının ChatGPT hakkındaki görüşlerinden bilgiye hızlı erişim, geri dönüt, yeterli düzeyde açıklayıcı bilgi sağlama, şekil ve görsel eksikliği, tutarsız bilgi vermesi ve matematik alanında eksik olması öne çıkmaktadır.

Anahtar Kelimeler: ChatGPT, ChatBot, OpenAI, AIEd, matematik eğitimi, yapay zeka

Atıf: Çiftçi, Z., & Öz, T., & Çiftçi O. (2026). Pre-Service Teachers' Experiences with Using Chatgpt for Mathematical Tasks. *İZÜ Eğitim Dergisi, 8*(1), 1–26. https://doi.org/10.46423/izujed.1793325

Sorumlu yazar: Tuba Öz tkaplan@atauni.edu.tr

Bu makale açık erişimli olup, CC BY-NC 4.0 lisansı ile dağıtılmaktadır.

1. INTRODUCTION

Artificial intelligence (AI) is the scientific and engineering field that develops systems capable of tasks requiring intelligence (Xu et al., 2021). In recent years, the rising impact of AI technologies has manifested itself in education and research as well as in many other fields. From the earliest levels of education, such as infancy, to the highest postgraduate levels, one of the main mechanisms by which AI has impacted education is individualized learning practices (Ocaña-Fernández et al., 2019). Furthermore, applying AI to natural language processing has led to the creation of intelligent chatbots that can understand and respond to human speech (Caldarini et al., 2022). The integration of artificial intelligence in learning environments has been more feasible because of the development of language models such as ChatGPT, which offers an interface that simulates human thought processes (Domenech, 2023). In its report on AI, the Contact North online learning community, one of Canada's largest non-profit organizations, stated that AI technology is inextricably linked to the future of higher education (Contact North, 2018, p. 5). The use of AI in higher education offers many opportunities and benefits that personalize learning and increase the effectiveness of the educational process and student motivation (Castrillón et al., 2020). The use of artificial intelligence in higher education can enhance learning outcomes by facilitating more engaging learning experiences, enhancing operational efficiency, and furnishing real-time, actionable insight into student performance (Muhie Woldie, 2020). Adopting technologies such as ChatGPT, a powerful AI-powered chatbot, in Higher Education institutions can positively impact various academic activities as it will make it easier to respond to individual student needs (Dempere et al., 2023). Though it is being utilized across various sectors, the potential risks of integrating ChatGPT into education have been controversial (Farrokhnia et al., 2023). Nevertheless, when innovative technologies are introduced into the educational sphere, these debates are inevitable as they challenge established practices and necessitate that educators adapt to the inherent advantages and disadvantages (Qadir, 2022).

1.1 Artificial Intelligence and AIEd

Although the concept of artificial intelligence has become more popular in recent years, it is not a new concept. Turing (1950), starting from the question "Can machines think?", defined codes used in machines that enable the machine to think like a human. McCarthy et al. (1955) furthered these scientific studies and coined the term 'artificial intelligence' (Cristianini, 2016; Crompton Burke, 2023; Tlili et al., 2023). Artificial intelligence (AI) is a technological system that has the ability to simulate human-like replies, which include reasoning, learning, synthesizing, and self-correcting processes (Popenici Kerr, 2017; Shukla Shubhendu Vijay, 2013). In recent years, the use of machine learning and neural network technologies has led to improvements in the definition of artificial intelligence (Wang, 2019). Generative artificial intelligence (GenAI), defined as a subset of artificial intelligence, encompasses a group of machine learning algorithms. Instead of analyzing and interpreting existing data, GenAI focuses on creating new data or content (McKinsey Consultant, 2023).

Today's artificial intelligence examples amaze with their superhuman computational skills and the knowledge they possess (Cope et al., 2021). In the light of these skills, rapidly developing artificial intelligence technology is used in many fields such as engineering,

banking, health, and space. In the field of education, although its impact has been limited compared to other fields or sectors, it has gained increasing momentum in recent years and has gained significant success in this sector (Murphy, 2019). The term Artificial Intelligence in Education (AIEd) is used to imply the integration of artificial intelligence technologies or applications into education by students and teachers (Hwang et al., 2020). Since the 1970s, the field of AIEd has exerted a growing influence on educational technology, prompting a shift in focus towards optimising the learning process and enhancing student achievement (Southgate et al., 2019). When the factors that make artificial intelligence more special and different from other technologies in education and training processes (Noe, 2009) are examined, personalization of the teaching process comes to the fore. In other words, facilitating personalized learning is among the goals of AIEd (Chen Liu, 2007; Hwang et al., 2014; Hwang et al., 2020). Artificial intelligence has unique benefits for individual users and can impact the tasks involved in the learning process in various ways (Kuprenko, 2020). Therefore, the potential of ChatGPT, one of the current applications of artificial intelligence, to be used in education and training is a new topic of curiosity.

1.2 ChatGPT and Education

Generative Pre-Trained Transducer (GPT) is a specific type of Generative AI model that employs deep learning techniques to facilitate the generation of natural language textual content (Chan, 2023). Many application tools have been created for generative AI that simulates the human brain by processing large amounts of data, such as Rytr, Jasper, Bing, ChatGPT, and CopyAI (Adeshola et al., 2023; Bengio et al., 2021; Zawacki-Richter et al., 2019). Farrokhnia et al. (2023) described that ChatGPT employs a more advanced language processing model than other chatbots, which is called GPT. Thus, ChatGPT has grown to be the fastest-developing user application in the world's history since its launch on November 30, 2022 (Tlili et al., 2023). ChatGPT, a generative AI-powered chatbot, was released by OpenAI, a community of researchers working on artificial intelligence. ChatGPT is furnished with a substantial language model, which endows it with the capacity to generate original text in response to prompts issued by users (OpenAI, 2022). The technology is operated either by subscription (ChatGPT 4.0) or free (ChatGPT 3.5) via an OpenAI account (OpenAI, 2022). ChatGPT is an essential tool for a variety of writing applications such as essays, marketing materials, brainstorming, coding tasks, and research studies (Dowling Lucey, 2023; Nautiyal et al., 2023; Owens, 2023). Generative AI can further facilitate to enhance students' learning experience by providing personalized, immediate feedback and adjusting to individual learning styles (Atlas, 2023; Chan Hu, 2023; Luckin, 2017). Today's students, called "millennials" or "digital natives", also tend to integrate these new technologies into their learning environments (Anega Alemu, 2023; Duh Dabula, 2021; Johri et al., 2014). Large language models such as ChatGPT have been well received among educators for their integration into different educational settings (Arif et al., 2023; Duong et al., 2023; Farrokhnia et al., 2023; Gilson et al., 2023; Kasneci et al., 2023; Lo, 2023; Mhlanga, 2023; Rospigliosi, 2023; Rudolph et al., 2023; Thurzo et al., 2023; Tlili et al., 2023). There is evidence that artificial intelligence is a valuable tool for solving multifaceted problems in numerous fields, such as education (Ouyang et al., 2022). The potential of using ChatGPT in the field of education is the subject of many current studies. In these studies, opportunities, advantages,

threats, and limitations in terms of the integration of ChatGPT in the field of education and training have been discussed (Alabool, 2023; Farrokhnia et al., 2023; Grassini, 2023; Loos et al., 2023; Michel-Villarreal et al., 2023; Zhu et al., 2023). According to Lo (2023), who recapitulated the potential advantages and issues being raised by experts, ChatGPT can help teachers with course material creation, idea generation, language translation, and testing; and students with questioning, summarizing, facilitating collaboration, checking concepts and preparing for an exam, outlining, and giving feedback. On the other hand, potential disadvantages of ChatGPT include accuracy, reliability, and plagiarism.

The possible impacts of ChatGPT in various fields (Adetayo, 2023; Dwivedi et al., 2023; Kasneci et al., 2023; Nautiyal et al., 2023; Siegerink et al., 2023), including higher education, have aroused great interest (Eke, 2023; Qi et al., 2023). Higher education institutions should make more efforts to see the benefits of artificial intelligence technologies and their effects on education and training processes (Muhie Woldie, 2020). Crompton Burke (2023), who examined artificial intelligence studies in higher education, emphasized that researchers in faculties of education have an increasing role; the importance of sharing pedagogical results with researchers in other disciplines, and that studies should become more widespread. In particular, Hwang and Tu (2021) mention that mathematics education researchers are of the background in artificial intelligence research in mathematics education and that they should use artificial intelligence technology in their research. In addition, they stated that AI applications in mathematics education are still in their infancy and researchers focus on using AI technologies to solve basic problems in mathematics courses. However, Pepin et al. (2025) identifies key themes, such as ChatGPT's capability to assist in understanding mathematical concepts, facilitate lesson planning, support assessment design, promote personalized learning, and foster collaboration. Additionally, there are numerous studies in the field of mathematics (Asare et al., 2023; Getenet, 2024; Frieder et al., 2023; Jia et al., 2024; Pavlova, 2024; Wardat, 2023; Zhou, 2022). In this context, the ability to use and apply artificial intelligence technologies effectively has become very important (Bozkurt, 2023). When it comes to the effective use of artificial intelligence technologies, prompt generation comes to mind. Commands that can accurately explain what and how to ask for when creating prompts will provide more accurate results, while unclear prompt open-ended residences will create more uncertain results (Liu Chilton, 2022). Because the slightest word change in the prompts can cause different outputs (Zhou et al., 2022). For all these reasons, as mathematics educators, we find it valuable to conduct research on artificial intelligence and its applications in our field. Accordingly, this study aims to determine how pre-service teachers create prompts when using ChatGPT to solve mathematics tasks, the accuracy of the answers of pre-service teachers and ChatGPT, and the opinions of pre-service teachers about ChatGPT. For the purpose of the study;

• How are the prompts created by pre-service teachers when using ChatGPT in mathematics tasks?

• How accurate are the answers given by ChatGPT and pre-service teachers to mathematics tasks?

• What are the opinions of pre-service teachers using ChatGPT about ChatGPT? answers to these questions were sought. It is thought that our research findings will contribute to the

expanding literature on the use of ChatGPT in the education and training process within the scope of AIEd.

2. METHOD

2.1 Research Design and Participants

In this study, in which the potential of using ChatGPT in mathematics education in higher education was investigated case study was preferred. Case study method is preferred when "why" and "why not" questions are asked and used in situations where researchers have no control over events or phenomena (Yin, 2009). Accordingly, the potential use of ChatGPT was examined by focusing on pre-service teachers' prompt construction processes, their interactions with ChatGPT, the accuracy of generated responses, and their subjective experiences and perceptions. This approach was preferred because it enables an in-depth and holistic investigation of both learning outcomes and underlying interaction processes. Furthermore, the case study design allowed the integration of multiple data sources, including task solutions, ChatGPT interaction screenshots, and opinion forms, thereby enhancing the richness and credibility of the findings through data triangulation. Consequently, the case study method was selected to provide a comprehensive understanding of how ChatGPT is used in mathematics tasks and how pre-service teachers experience this emerging educational technology. Although experiential data were included in the study, the primary analytical focus was placed on interaction processes, prompt construction patterns, and task-based performance outcomes. Therefore, the study was structured within a multiple-case case study framework to allow a comprehensive examination of these processes in an authentic learning context.

The research was conducted with 32 pre-service primary school teachers at a public university. The sample of this study consisted only of pre-service primary school teachers, and the sample size was relatively limited. In addition, the mathematical tasks used in the study were restricted to specific content areas. These factors may limit the generalizability of the findings to different educational levels, disciplines, and learning contexts.

2.2 Data Collection Tools

Data collection tools in the research comprised two distinct mathematical problems that the pre-service teachers were requested to solve with the use of ChatGPT, screenshots that captured the entirety of the conversations between the pre-service teachers and ChatGPT throughout the learning process, and a single-question opinion survey aimed at the pre-service teachers. Information about the data collection tools and the purposes of using the data collection tools are given in Table 1.

The data collection tools were structured to be directly aligned with the research questions. To address the first research question, which focuses on examining pre-service teachers' prompt construction processes and interaction patterns during ChatGPT use, ChatGPT interaction screenshots were selected because they enable direct observation of the interaction

Table 1. Data Collection Tools and The Purposes of Using Data Collection Tools

Data Collection Tool	Task	Purpose of Use in Research
Mathematical Tasks	1. Task: Find and prove a formula for this formula.	To examine the accuracy of ChatGPT-generated responses and pre-service teachers' solutions in the algebra domain
	2. Task: How do you draw the perpendicular bisector of a given line segment using a compass and an unmeasured ruler? Provide a step-by-step answer.	To examine the accuracy of ChatGPT-generated responses and pre-service teachers' solutions in the geometry domain and to evaluate ChatGPT's ability to provide procedural and visual-spatial explanations.
ChatGPT screenshots	-	Analysing ChatGPT answers and pre-service teachers' prompt construction processes
Question in the opinion survey	Evaluate your experience with the ChatGPT application	Setting opinions about ChatGPT

Note: All data collection tools were administered during the academic semester.

process. This approach made it possible to analyze user–ChatGPT interactions without relying solely on participants' self-reported statements.

The second research question aimed to determine the mathematical response accuracy of both ChatGPT and pre-service teachers. Therefore, mathematical tasks were used as the primary data source. In addition, ChatGPT interaction screenshots were analyzed as a supplementary data source, as they allowed ChatGPT-generated responses to be recorded in their original form and directly compared with pre-service teachers' solutions.

The third research question focused on revealing pre-service teachers' experiences and perceptions regarding ChatGPT use. For this purpose, an open-ended questionnaire was employed. This approach facilitated the in-depth and context-sensitive examination of user experiences related to emerging educational technologies.

To enhance the validity of the data collection tools, expert opinion was sought prior to the data collection process. During this process, the opinions of two experts in mathematics education were obtained. The experts evaluated the mathematical tasks and the open-ended opinion question in terms of content appropriateness, language clarity, and alignment with the research objectives. Necessary adjustments were made based on the experts' feedback.

2.3 Research Process

The implementation process of the research started by inviting pre-service teachers from the relevant department to participate voluntarily in the study. A total of 32 pre-service teachers (22 female and 10 male) who agreed to participate were included and asked to download the ChatGPT application to their mobile phones. Prior to the implementation, participants were verbally asked whether they had previously used ChatGPT or similar generative artificial intelligence tools. Only those who confirmed that they had no prior experience with ChatGPT were included in the sample. All participants had previously completed

an "Instructional Technologies" course during their first year of undergraduate education. However, this course did not include content related to artificial intelligence applications or prompt engineering practices. Therefore, participants did not have formal training in the pedagogical use of AI-based tools. After downloading the ChatGPT application, participants were seated in a traditional classroom arrangement. Since the pre-service teachers had no prior experience with ChatGPT, one of the researchers provided a short orientation session introducing the background, main functions, usage purposes, and basic operational features of the application. Participants were then given a brief free exploration period to familiarize themselves with the interface. During the implementation phase, participants were given a total of 30 minutes to use ChatGPT for both mathematical tasks collectively, rather than allocating separate time for each task. After completing the ChatGPT interaction phase, participants were instructed to close the application and were given an additional 20 minutes to independently write their final solutions on paper. Throughout the implementation process, it was ensured that participants did not use any additional resources other than ChatGPT. All ChatGPT interaction logs were collected by asking participants to send screenshots of their conversations to the researcher via email. Following the task completion, a one-item open-ended questionnaire was administered to collect participants' opinions regarding their experiences with using ChatGPT. All collected data were systematically classified and securely stored.

Mathematical tasks were prepared by the researchers before the implementation process. Before starting the application, information about ChatGPT was given by the researcher. In the data collection process, the implementation of the mathematical tasks, the organisation of the data afterwards, the preparation of the data for analysis and the analysis processes were carried out by the researchers.

For the validity and reliability of the study, expert opinion, data storage, detailed explanation (implementation stages, learning environment, data analysis method, and the role of the researcher), consensus in the data analysis process, and direct quotations were used. With regard to the ethical considerations of the study, the real names of the participants were not used. Verbal consent was sought from the pre-service teachers indicating their willingness to take part. The participants were verbally informed. Approval by the appropriate study-specific committee for research involving humans and/or animals. Ethics committee approval for this study was obtained from [Atatürk University] Ethics Committee with the decision dated [31.01.2024] and numbered [68].

This study was conducted using ChatGPT version 3.5; therefore, the findings may not fully reflect the performance of more recent model versions. Moreover, the free accessibility of the application may have influenced participants' interaction patterns and ease of use. In addition, noticeable differences were observed in ChatGPT's responses when the same geometry question was asked in Turkish and English. This indicates that the study is also subject to language-related limitations.

2.4 Data Analysis

The participants in the study were coded as S1, S2, …, S32. The answers given by pre-service teachers to each mathematical task were coded as "1" if they were completely

correct and "0" if they were incorrect or incomplete. To examine the accuracy of ChatGPT's responses to mathematical tasks, all ChatGPT interaction outputs were evaluated by three mathematics education experts. During this process, responses that were scientifically and mathematically correct were coded as "correct", whereas responses containing conceptual, procedural, or logical errors were coded as "incorrect". Qualitative data obtained from ChatGPT interaction screenshots related to prompt creation processes and pre-service teachers' experiences with ChatGPT were analyzed using content analysis. During the analysis of ChatGPT interaction screenshots, the screenshots belonging to each participant were first organized and systematically coded. Subsequently, the interaction texts, including user prompts and ChatGPT responses, were examined in detail. Meaningful data segments related to prompt generation strategies, interaction patterns, and user behaviors were identified and labeled. An Excel database was created to summarize each participant's ChatGPT interactions. This file included descriptive information such as the accuracy of ChatGPT's responses, the characteristics of the prompts used for each question, and the total number of questions asked by each participant. Codes were then generated inductively from the data. For example, when a pre-service teacher copied a mathematical task directly into ChatGPT without any modification, this interaction was coded as "asking the whole task". Requests for diagrams or visual explanations were coded as "asking for explanations with different representations". Similarly, prompts such as "What is the difference between the two formulas?" or "Why did we multiply by 2?" were grouped under the code "asking an interrogative question". During the coding process, researchers aimed to select the most representative expressions to ensure that the codes accurately reflected the underlying interaction patterns observed in the data. Similarly, participants' opinions about ChatGPT were coded by researchers and then categorized as positive or negative features.

A consensus-based coding approach was adopted throughout the qualitative analysis process. The researchers analyzed the data collaboratively, generated codes through joint discussions, and reached mutual agreement to finalize the coding. This approach was implemented to enhance analytical consistency and interpretive reliability. Furthermore, methodological triangulation was achieved through the combined use of mathematical task solutions, ChatGPT interaction screenshots, and opinion forms, which strengthened the credibility and reliability of the findings.

3. FINDINGS

3.1 Findings on How Pre-Service Teachers Use ChatGPT

The prompts that pre-service teachers directed to ChatGPT while solving the specified tasks were investigated. The prompts used by the pre-service teachers were named by the researchers as a result of the content analysis. The types of prompts used by pre-service teachers to ChatGPT while solving the tasks are presented in Table 2.

When Table 2 is analyzed, it is noteworthy that the pre-service teachers mostly created prompts by asking the whole task for the task solution in ChatGPT; they used 'asking a part of the task', 'asking an interrogative question', 'asking for an alternative solution' and 'feedback' prompts less frequently. On the other hand, the prompts 'asking for additional information'

Table 2. The Types of Prompts That Pre-Service Teachers Created in ChatGPT While Solving The Tasks

Prompt Type	Number of prompts for Task 1	How many different pre-service teachers used the prompt	Number of prompts for Task 2	How many different pre-service teachers used the prompt
Asking the whole task	23	23	33	31
Asking a part of the task	11	11	-	-
Asking the wrong question	1	1	1	1
Asking for an alternative solution	5	4	19	12
Asking an interrogative question	12	10	8	6
Feedback	2	2	2	2
Request additional information	2	2	-	-
Asking for explanations with different representations	1	1	2	2

Note: Task 1 and Task 2 represent the algebra and geometry domains respectively.

and 'asking for a part of the task', which appeared in Task 1, were not encountered in Task 2. The fact that Task 1 consists of two parts finding and proving formulae may be the reason for this code difference. The reason why the number of requests exceeds the number of people in Table 2 is that some students wrote more than one prompt. Examples of prompts created by pre-service teachers and related codes are presented below.

Asking a part of the task: "Can you find a formula for 1+2+3+4+5+6...+n=? " (S2)
Asking for an alternative solution: "What other method can I use besides this formula?" (S18)
Feedback: "I said that this addition can also be found as the product of the middle number and the last number, is that a wrong statement?" (S10)
Feedback: "Is drawing the middle a healthy way to achieve results?" (S18)
Asking for explanations with different representations: "Can you show the drawing of the center strut with a figure?" (S30)
Request additional information: "Can you prove more?" (S21)
Asking the wrong question: "Find and prove a formula for 1+2+3+4+4+n= ?" (S20)

3.2 Findings on the Analysis of the Responses of Pre-Service Teachers and ChatGPT to the Tasks

In this section, the outputs generated by ChatGPT as a result of pre-service teachers' interaction with ChatGPT for mathematics homework were examined in terms of being

scientifically correct or not. Within the scope of the research, the findings regarding the accuracy of the answers of all questions that pre-service teachers asked ChatGPT in their interactions with ChatGPT are presented in Table 3 as frequency and percentage.

Table 3. Frequency/Percentage Distribution of The Accuracy of ChatGPT's Answers to The Questions Asked

Tasks	Accuracy of ChatGPT's Answer	Frequency	Percentage
Task 1 (Algebra Field)	Correct	39	67%
	Incorrect	19	33%
	Total	58	100%
Task 2 (Geometry Field)	Correct	5	8%
	Incorrect	61	92%
	Total	66	100%

Note: Percentages are rounded to the nearest whole number.

When Table 3 is analyzed, it is seen that all of the pre-service teachers asked each given question to ChatGPT. Apart from this, some pre-service teachers asked more than one question to ChatGPT, such as asking for an alternative solution, questioning the answer, and elaborating. For this reason, when Table 3 is analyzed, it is seen that the total number of questions is higher than the number of pre-service teachers. When the accuracy rates of ChatGPT's answers to the questions are analyzed, it is seen that 67% of the answers to the algebra question were correct, while 92% of the answers to the geometry question were incorrect. When the answers given to both questions are considered together, it is seen that ChatGPT gave 65% incorrect answers. These incorrect pre-service answers are due to mathematical inaccuracies, language and notation errors, contradictory and inconsistent explanations. Artificial intelligence and AIEd Figure 1-5 provides examples of the specified answers of ChatGPT. In addition to these findings, Table 4 includes the answers of each pre-service teacher working on mathematical tasks with the help of ChatGPT and the accuracy values of ChatGPT's answers on a pre-service teacher basis in detail. Thus, it is aimed to provide a more detailed description.

When Table 4 is analyzed, it is seen that 10 pre-service teachers (31%) gave correct answers to the 1st task, but only one pre-service teacher (3%) gave correct answers to the 2nd task. It is also noteworthy that some of the pre-service teachers (S5, S7, S8, S14, S15, ...) gave incorrect answers to the mathematical tasks, although ChatGPT gave correct answers to the pre-service teachers' prompts. Another notable finding is that the rate of incorrect answers generated by ChatGPT for Task 2 (92%) is very high, whereas the rate of correct answers provided by pre-service teachers for this task is very low (3%). This indicates that the correctness of the answers generated by ChatGPT can influence the accuracy of the user's answers. A few sample screenshots of the answers generated by ChatGPT are provided below.

In Figure 1, ChatGPT provided a correct answer to the question asked by S1 in algebra with clear step-by-step explanations using the relevant mathematical notations. However, when the same question was asked by S6, ChatGPT gave an incorrect answer by using unequal expressions as seen in Figure 2.

Table 4. Accuracy of Pre-Service Teachers' and ChatGPT's Answers to Math Tasks

	Task 1			Task 2		
	Accuracy value of the pre-service teacher's answer	*Accuracy value of ChatGPT's response*		*Accuracy value of the pre-service teacher's answer*	*Accuracy value of ChatGPT's response*	
Student	*True/False*	True	False	*True/False*	True	False
S1	0	1	0	0	0	1
S2	0	2	1	0	0	2
S3	1	1	0	0	0	1
S4	0	0	2	0	1	3
S5	0	1	0	0	1	0
S6	0	0	2	0	0	2
S7	0	3	0	0	0	1
S8	0	3	0	0	0	3
S9	1	1	0	0	0	5
S10	0	1	1	0	0	2
S11	0	3	0	0	0	3
S12	0	0	1	0	0	1
S13	1	1	1	0	1	0
S14	0	1	0	0	0	1
S15	0	1	0	0	0	2
S16	0	2	0	0	0	2
S17	0	1	0	0	0	1
S18	0	2	2	0	0	4
S19	1	0	1	0	0	1
S20	0	0	1	1	0	1
S21	0	2	0	0	0	5
S22	0	0	1	0	0	1
S23	1	2	0	0	0	3
S24	0	0	1	0	0	1
S25	0	0	1	0	0	1
S26	0	1	1	0	0	1
S27	1	4	0	0	0	5
S28	0	1	1	0	0	2
S29	0	1	0	0	1	1
S30	0	2	0	0	0	2
S31	0	1	1	0	0	2
S32	0	1	1	0	1	1
Total	31%	39 (67%)	19 (33%)	3%	5 (8%)	61 (92%)

Note: True/False column represents the accuracy value of the pre-service teacher's answer, while True and False columns represent the accuracy values of ChatGPT's responses.

When the ChatGPT answer in Figure 2 is analyzed, it is seen that ChatGPT takes the sum of n n+1 terms as n+1 instead of the correct expression n(n+1); in addition, it takes the

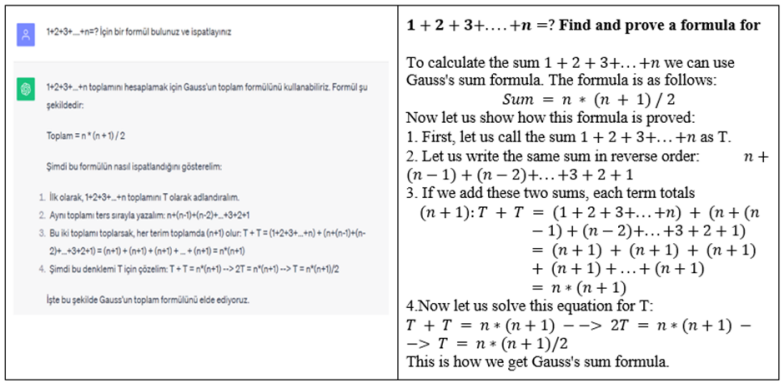

Figure 1. ChatGPT's Correct Answer to Algebra Task and its English Equivalent

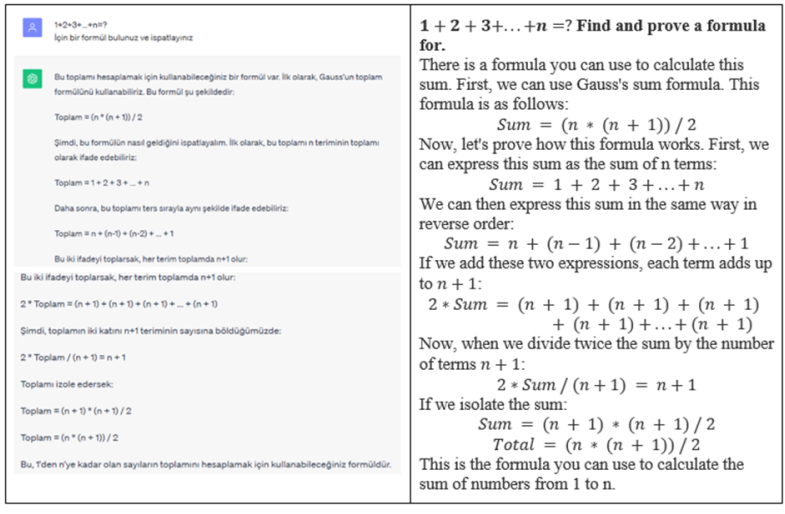

Figure 2. ChatGPT's Incorrect Answer to The Algebra Task and its English Equivalent

product of (n+1)*(n+1) as equal to n*(n+1). These equations are mathematically incorrect. In other words, ChatGPT has shown the wrong steps as correct to reach the formula it has determined. S4 asked ChatGPT for a simpler solution for the same question. The answer of ChatGPT, which offers a simpler solution in Figure 3, was also included in the incorrect category.

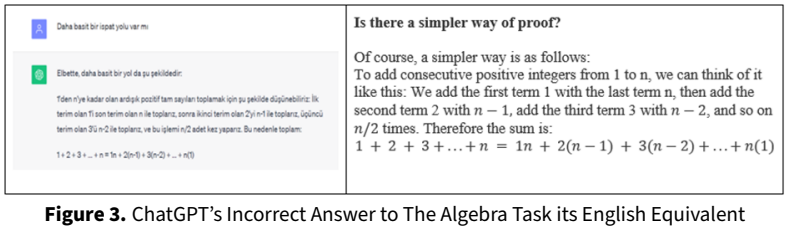

Figure 3. ChatGPT's Incorrect Answer to The Algebra Task its English Equivalent

When the ChatGPT answer in Figure 3 is examined, it is seen that although he stated that he would use the addition operation in the solution process, he used the multiplication operation while performing the operation. Here, both a contradictory answer was given and the operation was not mathematically correct. ChatGPT screenshots of the answers to the question in the geometry domain are given below.

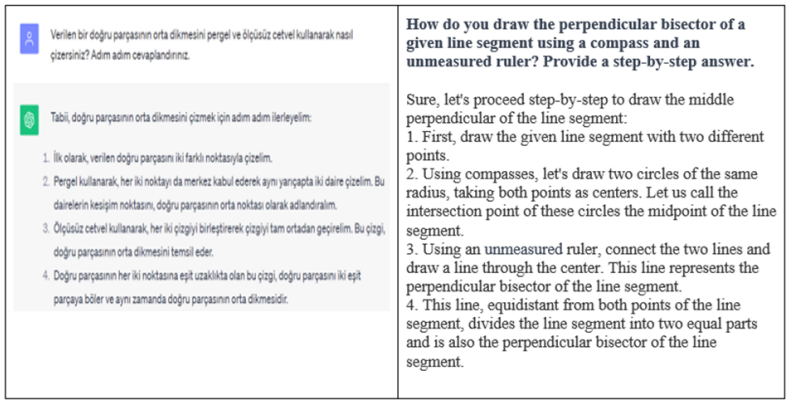

Figure 4. ChatGPT's Correct Answer to The Geometry Task and its English Equivalent

In Figure 4, it is seen that ChatGPT explained step-by-step, accurately, and clearly how to find the midpoint of a line using a compass and an unmeasured ruler to the question asked by S5 in geometry. S26 asked the same question to ChatGPT, but ChatGPT gave the wrong answer as seen in Figure 5.

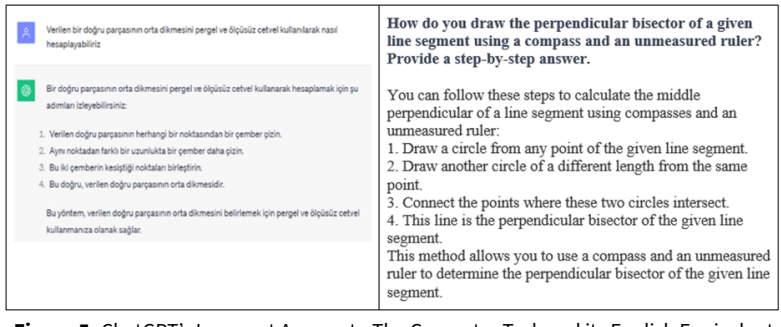

Figure 5. ChatGPT's Incorrect Answer to The Geometry Task and its English Equivalent

When Figure 5 is analyzed, ChatGPT explained in unclear steps how to find the midpoint of a line using a compass and an unmeasured ruler. ChatGPT had the students draw circles on the line segment with randomly selected centers. Then he drew the line segment formed by the union of the points where the circles intersect. Finally, he stated that this line segment is the perpendicular bisector of the line segment. These steps are incorrect for finding the perpendicular bisector of a line segment.

3.3 Opinions of Pre-Service Teachers Using ChatGPT About Chatgpt

The opinions of the pre-service teachers about the ChatGPT application were examined and positive/negative opinion categories about ChatGPT were created. Figure 1 shows these two categories, codes, and their frequencies. The discrepancy between the total frequency values of the codes in Figure 6 and the actual number of pre-service teachers attributed to the fact that these individuals articulated multiple perspectives.

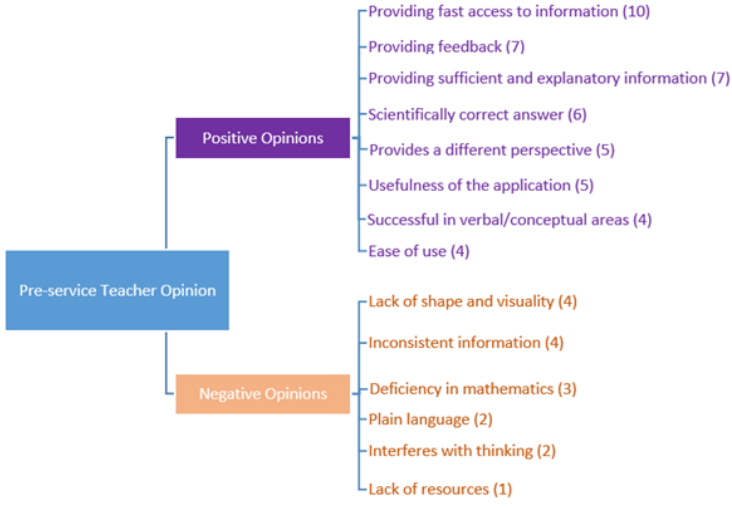

Figure 6. Positive and Negative Aspects of The ChatGPT Application

The majority of the pre-service teachers who participated in the application expressed positive opinions about the ChatGPT application. Among these opinions, quick access to information, and providing sufficient/explanatory information and feedback were the most frequently mentioned positive features. In addition, the least mentioned positive feature is the ease of use of the application. Some of the pre-service teachers' statements about these positive features and the supporting ChatGPT images are given below.

> *"I found the application useful because it provides us with more accurate information quickly without getting lost among all the information on the internet."(S27)*

> *"I like the application because it helps me with questions and proofs that I do not know the answer to, with good explanations that are sufficient and satisfactory in research... (S19).*

> *"I liked the application. After using the application, I got more answers to the questions that I could not fully prove in the first test or that I wanted to prove but got stuck at some point. I got a very sufficient explanation by asking the points that did not satisfy me or that I did not understand in the answer given." (S18)*

Figure 7 and Figure 8 show that the pre-service teacher received feedback from the application about the parts she did not understand with the question "Why... ?" after the answer she received from ChatGPT.

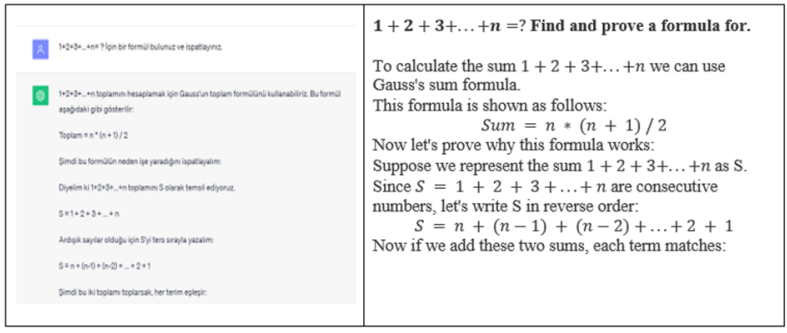

Figure 7. An Excerpt from S19's ChatGPT Correspondence Image and its English Equivalent

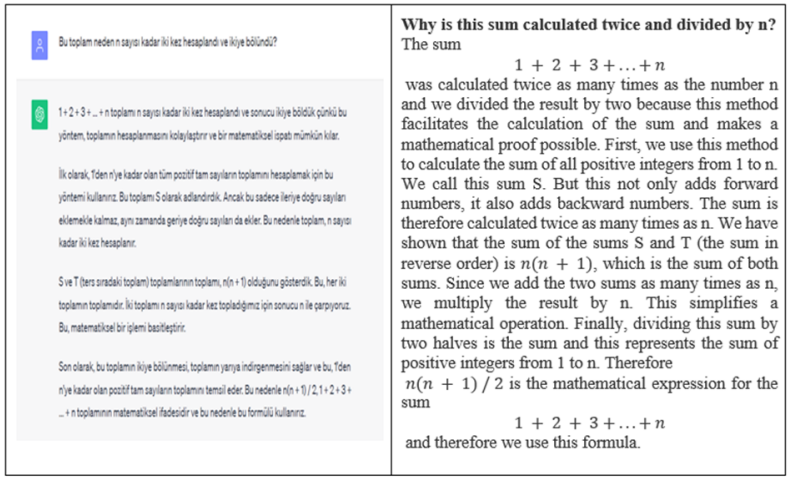

Figure 8. An Excerpt From S18's ChatGPT Correspondence Image and its English Equivalent

"I liked the application very much based on verbal courses... It is an application that I will use constantly for verbal knowledge." (S13)

Some of the pre-service teachers who participated in the implementation also mentioned the negative aspects of the ChatGPT application. The most frequently mentioned negative aspects are that the application lacks figures and visuals, presents inconsistent information and is insufficient in mathematics. The fact that the application does not provide resources and its language is not simple are among the least frequently mentioned negative opinions. Some pre-service teachers' statements about negative opinions and supporting ChatGPT visuals are given below.

"I don't like the application because it doesn't have any visual aid. It is very difficult to solve geometry questions with written information..." (S12)

Similarly, as seen in Figure 9, S9's question to ChatGPT, "Can you explain this question to me with a visual?" supports the opinion that the answers given by the application lacked shapes and visuals.

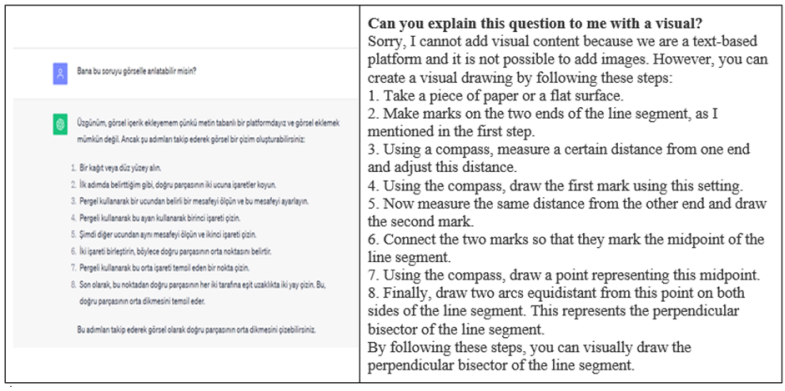

Figure 9. An Excerpt From S9's ChatGPT Correspondence Image and Its English Equivalent

I like the application. We can access the information we are curious about. But how accurate is this information? I have concerns about trust as we cannot be sure about this." (S31).

In Figure 10, when S31 questioned ChatGPT's answer, ChatGPT gave a wrong answer. It is thought that S31's concern about trust may stem from this reason.

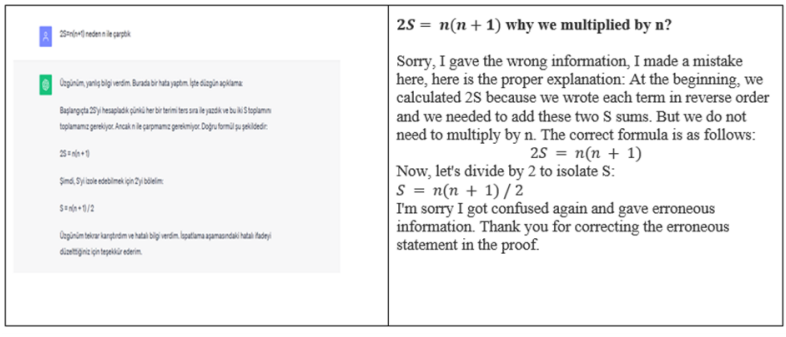

Figure 10. An Excerpt from S13's ChatGPT Correspondence Image and its English Equivalent

"I had difficulty in understanding exactly where the given operations and equations came from and how he found the equations. For this reason, I do not think it is an application that I can use in the mathematics course." (S13)

" ... only the fact that it does not specify the source can leave the application one click behind." (S11)

"...I think the Turkish translation of the application is not very good. The answers are confusing and inconsistent." (S5)

4. CONCLUSION AND DISCUSSION

This study aimed to explore the potential of utilizing ChatGPT in mathematics education by investigating three key aspects: the ways in which pre-service teachers formulate prompts

while using ChatGPT to solve mathematical tasks, the accuracy of the responses provided by both pre-service teachers and ChatGPT in addressing mathematical problems, and the perceptions of pre-service teachers regarding the use of ChatGPT.

The results of the study indicated that pre-service teachers who engaged with mathematical tasks using ChatGPT generated a variety of prompts. It was concluded that the majority of the prompts generated by the pre-service teachers primarily involved simply restating the entirety of the mathematical task. However, chatbots can produce more appropriate and original answers depending on the nature of the prompts. Tapan Brotin (2023) too emphasises that purposeful and well-expressed commands directed to artificial intelligence will ensure the production of answers with similar characteristics. In ChatGPT, complex tasks tend to have higher error rates than simple tasks (OpenAI, 2024). The fact that the pre-service teachers asked the ChatGPT whole of the given mathematical task without dividing it into simpler sub-questions suggests that pre-service teachers had limited proficiency in using the chatbot effectively. The reason for this situation may be that pre-service teachers are not competent and experienced in creating prompts. In addition, the nature of the mathematical tasks also caused some differences in the prompts. As a matter of fact, in Task 2, which was asked from the field of geometry, since ChatGPT provided more verbal answers, the students did not understand the answer given and tended to ask the chatbot for visual or figure prompts. Similarly, Vardat et al., (2023) too emphasised that ChatGPT lacks understanding of geometry and this situation makes it lag behind in correcting misconceptions. Apart from asking the whole task, it was observed that the pre-service teachers used some different prompts although they were not guided by the researchers. Among these prompts, asking a part of the task, feedback, and asking an interrogative question can be effective prompts to activate the chatbot. The fact that the pre-service teachers stated that they did not receive any training on creating prompts in ChatGPT and that their frequency of use was low suggests that the pre-service teachers have the ability to create these prompts spontaneously, albeit to a lesser extent. This means that it is possible to obtain more effective and accurate results if users are supported in creating prompts. Similar results are in parallel with Noster et al. (2024). As emphasized by Khurma et al. (2023, 2024), human–AI interaction and collaboration are crucial for obtaining the most accurate and satisfactory answers for educators and students. This interaction enables users to continuously refine their prompts until they achieve the most effective responses, which, in turn, indirectly enhances their ability to assess the efficiency of their prompts. Based on these findings, future studies should focus on designing instructional interventions that explicitly teach pre-service teachers how to formulate effective prompts for educational purposes. In particular, training activities that emphasize task decomposition, step-by-step questioning strategies, and reflective prompt refinement may help users interact with ChatGPT more productively. Furthermore, teacher education programs may consider integrating AI literacy and prompt engineering components into their curricula in order to prepare future teachers for effective human–AI collaboration. Such integration may contribute to the development of pedagogical competencies that enable teachers to critically evaluate AI-generated responses rather than using them passively. Therefore, future research may investigate the impact of systematic prompt training on students' mathematical reasoning performance and their ability to evaluate the reliability of AI-generated solutions.

When the accuracy of the answers given by ChatGPT and pre-service teachers to the mathematical tasks were analyzed; pre-service teachers answered 67% correctly for task 1 and 3% correctly for Task 2 despite receiving ChatGPT help. Several factors may account for this situation, including the pre-service teachers' challenges in formulating effective prompts, the perceived quality of the responses provided by ChatGPT to mathematical questions, the fact that many pre-service teachers were encountering ChatGPT for the first time and had limited experience using it, their lack of trust in the accuracy of ChatGPT's answers, and instances of unsatisfactory responses from ChatGPT. ChatGPT presented different answers to each pre-service teacher. This can be an advantage for learning if the information presented is correct. However, within the scope of the mathematical tasks used in this study, ChatGPT produced incorrect responses in approximately 65% of the cases. Furthermore, a higher proportion of incorrect responses was observed in the geometry task (92%) compared to the algebra task (33%). It should be noted that these findings are based on a limited number of task types and should be interpreted as task-specific patterns rather than generalized performance indicators of ChatGPT across all mathematical domains. Although artificial intelligence systems act as personal tutors, they can sometimes give incorrect results for more specific situations because they are based on big data analysis (Chassignol et al., 2018). One of the primary limitations of ChatGPT remains its ability to provide accurate or complete answers to complex questions (Alabool, 2023). The complexity of the equation, the input data, and the prompts provided to ChatGPT significantly influence the efficiency and accuracy of its solutions. Nonetheless, it is anticipated that ChatGPT will demonstrate greater efficiency in addressing progressively complex mathematical problems (Vardat et al., 2023). The limitations of artificial intelligence are highlighted by the inaccuracies present in the information provided, underscoring the continued necessity for human oversight (Domenech, 2023). It is necessary to emphasize that learning only with ChatGPT is not effective by attributing a teacher role to ChatGPT in the education and training process. Because, when pre-service teachers ask a question to ChatGPT on a subject they do not know, they may have difficulty in analyzing the answer given because they do not know whether the answer is correct or not, in this sense, ChatGPT alone is not sufficient for teaching. In this direction, Yifan et al. (2023) mention that ChatGPT should be used as a tool to help students solve problems and gain knowledge rather than replacing the instructor in teaching. Mrabet Studholme (2023) also suggest that ChatGPT can be included in the educational environment as a complementary tool to aid research, clarify concepts, and provide additional practice. This is because, unlike search engines that provide a breakdown of individual links to the information sought, ChatGPT is an interactive conversational system that develops and expands answers and responds to challenges posed by the user (Rospigliosi, 2023). ChatGPT is interactive with the user, as it can link between the user's initial and subsequent questions. The opportunity for interaction in ChatGPT is important for users to diversify their questions and identify points that they do not understand. When this situation is used effectively, the answers given by ChatGPT can be handled critically and efforts can be made to reach the right information. Based on these findings, future instructional designs should incorporate structured guidance that teaches students how to verify AI-generated mathematical solutions. Moreover, the substantially higher error rate observed in geometry tasks suggests that domain-specific support mechanisms should be considered. Future studies may investigate the integration of ChatGPT with dynamic geometry software

or symbolic mathematics tools in order to improve performance in visually and conceptually complex mathematical domains. This hybrid approach may help overcome current limitations related to spatial reasoning and proof explanation. Furthermore, an additional language-related issue was observed during the implementation process. When the geometry task was posed in English, ChatGPT generated responses that differed substantially in clarity and procedural explanation quality compared to those produced in Turkish. This finding suggests that the linguistic formulation of prompts may directly influence the quality and structure of AI-generated mathematical explanations. Although ChatGPT officially supports multiple languages, variations in response quality across languages indicate that multilingual performance inconsistencies remain an important research issue. Therefore, future studies should systematically examine how language choice affects the accuracy, coherence, and pedagogical usefulness of AI-generated mathematical solutions across different linguistic contexts. In addition, classroom-based experimental studies comparing guided and unguided ChatGPT use may provide valuable insights into how teacher scaffolding influences students' ability to detect errors and interpret AI-generated feedback. Such findings may contribute to the development of pedagogical models in which ChatGPT is positioned not as a replacement for instructors, but as a supervised and complementary instructional tool.

Pre-service teachers think that ChatGPT is an application that provides quick access to information, sufficient explanatory information, and ease of use. In the study of Yifan et al. (2023), students stated that the interface of the application was simple and provided quick access to information. Similarly, lecturers in Firaina and Sulisworo's (2023) study and students in Essel et al.'s (2022) study stated that they were satisfied with the fast feedback of chatbots. ChatGPT offers the opportunity to access information quickly (Alabool, 2023). Therefore, it saves time and allows students to spend more time criticizing and evaluating information (Farrokhnia et al., 2023). In addition, pre-service teachers think that ChatGPT is a successful and useful application that provides feedback, gives scientifically correct answers, and is successful verbally and conceptually. These features of ChatGPT can be considered among the factors that facilitate the user's learning. Feedback, especially in the teaching process, requires a very short interaction cycle where the teacher has only seconds for feedback (Wiliam Thompson, 2017). ChatGPT has been adapted to meet the need to provide customized feedback and content by examining students' learning styles, preferences, and needs (Rawas, 2023). As a powerful tool for feedback, ChatGPT is useful in quickly identifying and correcting errors, thus providing a better understanding of the subject matter (Alabool, 2023). ChatGPT can provide instant feedback to students, facilitating their understanding and enabling them to learn from their mistakes. This immediate feedback promotes self-assessment and allows students to promptly address and correct their misconceptions (Taani Alabidi, 2024). Thanks to this feature of ChatGPT, students can use this application as an aid in their learning process (Dempere et al., 2023; Loos et al., 2023).

In addition to the positive aspects of ChatGPT, Aruleba et al. (2023) and Kasneci et al. (2023) mentioned the possibility that knowledge produced without effort may increase laziness and reduce critical thinking, problem-solving skills, and the desire to do their own research. In parallel with Kasneci et al. (2023), the majority of the pre-service teachers in our study reported that they completed their learning with ChatGPT within the first 5 minutes. It is clear that this time is not enough to make sense of the answers to the two questions.

As it is understood from this, students did not focus on learning because they tried to get answers from ChatGPT in a superficial way without forming a critical perspective. This situation was also included among the negative opinions of the pre-service teachers with the code of ChatGPT being an obstacle to thinking. Some of the negative opinions expressed by the pre-service teachers were that the answers were confusing and inconsistent and that the source of the information was not specified. Most of the pre-service teachers stated that they could not trust the accuracy of the data, so the application could be used to get an idea and have preliminary knowledge on a certain subject, although not on every subject. Similarly, Choudhury and Shamszare (2023) emphasized in their study that trust is very important for users to adopt ChatGPT. Therefore, inaccuracies in ChatGPT's answers, which may pose a problem in terms of academic rigor Lo (2023) and potential honesty issues reduce users' willingness to use the application (Essel et al., (2022)). Another negative opinion of the application is that ChatGPT lacks shape and visualization. This opinion is in line with Yifan et al. (2023) and Zhai (2022) who concluded that ChatGPT can be made more interesting by adding some interesting features, and interactive or game elements. Some pre-service teachers think that ChatGPT may be sufficient for verbal and interpretation questions, but it is not sufficient for mathematics and its language is not simple. This may be due to the way mathematical questions are asked and the use of symbols. Similarly, Domenech (2023) argues that ChatGPT is well suited for conveying knowledge in the humanities or social sciences where plain language is used, but its effectiveness may be limited in fields that rely heavily on mathematical and specialized notations due to its limited understanding of them. Frieder et al. (2023) also argue that, contrary to the many positive reports about the test-solving abilities of GPT-4 and ChatGPT, the overall mathematics performance of the applications was significantly below the level of a graduate student. Based on these findings, ChatGPT should be positioned in educational settings not as a direct information provider but as a guided support tool that facilitates learning processes. In order to prevent superficial learning resulting from students' tendency to seek rapid answers, it is important to integrate critical questioning, source verification, and justification-based explanation activities into teacher-guided classroom practices involving ChatGPT.

Furthermore, considering students' perceptions regarding the lack of visualization and mathematical expression support, the design of integrated learning environments in which ChatGPT is used together with dynamic geometry software, graphing tools, and interactive visual platforms is recommended. Such environments may contribute to the concretization of abstract mathematical concepts and support deeper conceptual understanding.

In addition, increasing AI literacy training within teacher education programs may enhance pre-service teachers' ability to critically evaluate the accuracy of AI-generated outputs and guide students more effectively. In this way, while preserving the advantages of ChatGPT in terms of rapid feedback and accessibility, potential risks related to surface-level learning and overreliance on AI tools can be pedagogically balanced.

Although ChatGPT demonstrated potential to support higher education students in solving mathematical tasks, several limitations were observed within the scope of this study. In particular, inaccuracies were identified in multi-step algebraic reasoning processes and in the justification stages of mathematical proofs, where the model occasionally produced

incomplete or logically inconsistent explanations. These findings indicate that current large language models require further development in domain-specific reasoning mechanisms, especially for symbolic manipulation, proof generation, and step-by-step mathematical verification processes.

In addition, since ChatGPT is primarily trained on English-based datasets, limitations related to mathematical terminology and conceptual clarity were observed when prompts were provided in Turkish. This suggests that multilingual training strategies and localized language models should be further developed to improve the effectiveness of AI-supported mathematics instruction in non-English educational contexts.

Conflict of Interest

There are no conflicts of interest among the authors.

Funding

No financial support was received from any institution or organization for this study.

Ethical Statement

The study was approved by the Ethics Committee of Atatürk University with decision number 01/68.

Author Contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by [all author] .The first draft of the manuscript was written by [all author] and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

ORCID

Zeynep Çiftçi https://orcid.org/0000-0002-3828-6230

Tuba Öz https://orcid.org/0000-0003-0536-9360

Orhan Çiftçi https://orcid.org/0000-0002-5969-9464

Declaration of Generative AI Use

During the preparation of this manuscript, AI tools (ChatGPT 3.5, Grammarly, DeepL) were used solely for language and grammar correction. No AI tools were used for idea generation, data analysis, literature review, or image creation, and none were listed as authors.

Acknowledgments

We would like to thank the pre-service teachers who participated in this study for their contributions.

References

Adeshola, I., & Adepoju, A. P. (2023). The opportunities and challenges of chatGPT in education. *Interactive Learning Environments*, 1–14. https://doi.org/10.1080/10494820.2023.2253858.

Adetayo, A. J. (2023). Artificial intelligence chatbots in academic libraries: The rise of ChatGPT. *Library Hi Tech News*, *40*(3). https://doi.org/10.1108/LHTN-01-2023-0007.

Alabool, H. M. (2023). ChatGPT in Education: SWOT analysis approach. *2023 International Conference on Information Technology (ICIT)*, 184–189. https://ieeexplore.ieee.org/document/10225801.

Anega, N. M., & Alemu, B. (2023). The effect of rural roads on consumption in Ethiopia. *Journal of Economics and Development*, *25*(3). https://doi.org/10.1108/JED-10-2022-0213.

Arif, T. B., Munaf, U., & Ul-Haque, I. (2023). The future of medical education and research: Is ChatGPT a blessing or blight in disguise? *Medical Education Online*, *28*(1), 2181052. https://doi.org/10.1080/10872981.2023.2181052.

Aruleba, K., Sanusi, I. T., Obaido, G., & Ogbuokiri, B. (2023). Integrating ChatGPT in a computer science course: Students perceptions and suggestions. *arXiv preprint*, *arXiv:2402.01640*, 1–13. https://doi.org/10.48550/arXiv.2402.01640.

Asare, B., Arthur, Y. D., & Boateng, F. O. (2023). Exploring the impact of ChatGPT on mathematics performance: The influential role of student interest. *Education Science and Management*, *1*(3), 158–168. https://doi.org/10.56578/esm010304.

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

Bengio, Y., Lecun, Y., & Hinton, G. (2021). Deep learning for AI. *Communications of the ACM*, *64*(7), 58–65. https://doi.org/10.1145/3448250.

Bozkurt, A. (2023). ChatGPT, generative artificial intelligence and algorithmic paradigm shift. *Journal Critical Reviews in Educational Sciences*, *4*(1), 63–72. https://doi.org/10.59320/alanyazin.1283282.

Caldarini, G., Jaf, S., & McGarry, K. (2022). A literature survey of recent advances in chatbots. *Information*, *13*(41), 2–22. https://doi.org/10.3390/info13010041.

Castrillón, O. D., Sarache, W., & Ruiz-Herrera, S. (2020). Prediction of academic performance using artificial intelligence techniques. *University Education*, *13*(1), 93–102. https://doi.org/10.4067/s0718-50062020000100093.

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

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *ArXiv preprint*, *arXiv:2305.00290*. https://doi.org/10.48550/arXiv.2305.00290.

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, C. J., & Liu, P. L. (2007). Personalized computer-assisted mathematics problem-solving program and its impact on Taiwanese students. *Journal of Computers in Mathematics and Science Teaching*, *26*(2), 105–121.

Choudhury, A., & Shamszare, H. (2023). Investigating the impact of user trust on the adoption and use of ChatGPT: Survey analysis. *Journal of Medical Internet Research*, *25*. https://doi.org/10.2196/47184.

Contact North. (2018). *Ten facts about artificial intelligence in teaching and learning*. https://teachonline.ca/sites/default/files/tools-trends/downloads/ten_facts_about_artificial_intelligence.pdf.

Cope, B., Kalantzis, M., & Searsmith, D. (2021). Artificial intelligence for education: Knowledge and its assessment in AI-enabled learning ecologies. *Educational Philosophy and Theory*, *53*(12), 1229–1245. https://doi.org/10.1080/00131857.2020.1728732.

Cristianini, N. (2016). Intelligence reinvented. *New Scientist*, *232*(3097), 37–41. https://doi.org/10.1016/S0262-4079(16)31992-3.

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), 1–22. https://doi.org/10.1186/s41239-023-00392-8.

Dempere, J., Modugu, K. P., Hesham, A., & Ramasamy, L. (2023). The impact of ChatGPT on higher education. *Frontiers in Education*, *8*. https://doi.org/10.3389/feduc.2023.1206936.

Domenech, J. (2023). ChatGPT in the classroom: Friend or foe? *9th International Conference on Higher Education Advances (HEAd'23)*, 339–347. https://doi.org/10.4995/HEAd23.2023.16179.

Dowling, M., & Lucey, B. (2023). ChatGPT for (finance) research: The Bananarama conjecture. *Finance Research Letters*, *53*, 103662. https://doi.org/10.1016/j.frl.2023.103662.

Duh, H. I., & Dabula, N. (2021). Millennials' socio-psychology and blood donation intention developed from social media communications: A survey of university students. *Telematics and Informatics*, *58*. https://doi.org/10.1016/j.tele.2020.101534.

Duong, C. D., Bui, D. T., Pham, H. T., & Vu, A. T. (2023). How effort expectancy and performance expectancy interact to trigger higher education students' uses of ChatGPT for learning. *Interactive Technology and Smart Education*, *ahead-of-print*. https://doi.org/10.1108/ITSE-05-2023-0096.

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). 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*, *53*. https://doi.org/10.1016/j.ijinfomgt.2023.102642.

Eke, D. O. (2023). ChatGPT and the rise of generative AI: Threat to academic integrity? *Journal of Responsible Technology*, *13*. https://doi.org/10.1016/j.jrt.2023.100060.

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), 1–19. 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.

Frieder, S., Pinchetti, L., Griffiths, R. R., Salvatori, T., Lukasiewicz, T., Petersen, P. C., Chevalier, A., & Berner, J. (2023). Mathematical capabilities of ChatGPT. *ArXiv preprint*, *arXiv:2301.13867*. https://doi.org/10.48550/arXiv.2301.13867.

Getenet, S. (2024). Pre-Service Teachers and ChatGPT in Multistrategy Problem-Solving: Implications for Mathematics Teaching in Primary Schools. *International Electronic Journal of Mathematics Education*, *19*(1). https://doi.org/10.29333/iejme/14141.

Gilson, A., Safranek, C. W., Huang, T., Socrates, V., Chi, L., Taylor, R. A., & Chartash, D. (2023). How does CHATGPT perform on the United States Medical Licensing Examination? The implications of large language models for medical education and knowledge assessment. *JMIR Medical Education*, *9*(1). https://doi.org/10.2196/45312.

Grassini, S. (2023). Shaping the future of education: Exploring the potential and consequences of AI and ChatGPT in educational settings. *Education Sciences*, *13*(7), 692. https://doi.org/10.3390/educsci13070692.

Hwang, G. J., Hung, P. H., Chen, N. S., & Liu, G. Z. (2014). Mindtool-assisted in-field learning (MAIL): An advanced ubiquitous learning project in Taiwan. *Journal of Educational Technology & Society*, *17*(2), 4–16.

Hwang, G. J., & Tu, Y. F. (2021). Roles and research trends of artificial intelligence in mathematics education: A bibliometric mapping analysis and systematic review. *Mathematics*, *9*(6), 584. https://doi.org/10.3390/math9060584.

Hwang, G. J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence*, *1*, 100001. https://doi.org/10.1016/j.caeai.2020.100001.

Jia, J., Wang, T., Zhang, Y., & Wang, G. (2024). The comparison of general tips for mathematical problem solving generated by generative AI with those generated by human teachers. *Asia Pacific Journal of Education*, *44*(1), 8–28. https://doi.org/10.1080/02188791.2023.2286920.

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274.

Khurma, O. A., Albahti, F., Ali, N., & Bustanji, A. (2024). AI ChatGPT and student engagement: Unraveling dimensions through PRISMA analysis for enhanced learning experiences. *Contemporary Educational Technology*, *16*(2). https://doi.org/10.30935/cedtech/14334.

Khurma, O. A., Ali, N., & Hashem, R. (2023). Critical reflections on CHATGPT in UAE education: Navigating equity and governance for safe and effective use. *International Journal of Emerging Technologies in Learning*, *18*(14), 188–199. https://doi.org/10.3991/ijet.v18i14.40935.

Kuprenko, V. (2020). *Artificial intelligence in education: Benefits, challenges, and use cases*. https://pub.towardsai.net/artificial-intelligence-in-education-benefits-challenges-and-use-cases-db52d8921f7a.

Liu, V., & Chilton, L. B. (2022). Design guidelines for prompt engineering text-to-image generative models. *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, 1–23. https://doi.org/10.1145/3491102.3501825.

Luckin, R. (2017). Towards artificial intelligence-based assessment systems. *Nature Human Behaviour*, *1*, 0028. https://doi.org/10.1038/s41562-016-0028.

McKinsey Consultant. (2023). *What is generative AI?* https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai.

Mhlanga, D. (2023). *Open AI in education, the responsible and ethical use of ChatGPT towards lifelong learning*. http://dx.doi.org/10.2139/ssrn.4354422.

Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D. E., Thierry-Aguilera, R., & Gerardou, F. S. (2023). Challenges and opportunities of generative AI for higher education as explained by ChatGPT. *Education Sciences*, *13*(9), 856. https://doi.org/10.3390/educsci13090856.

Mrabet, J., & Studholme, R. (2023). ChatGPT: A friend or a foe? *2023 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE)*, 269–274. https://ieeexplore.ieee.org/document/10131713.

Muhie, Y. A., & Woldie, A. B. (2020). Integration of artificial intelligence technologies in teaching and learning in higher education. *Science and Technology*, *10*(1), 1–7. https://doi.org/10.5923/j.scit.202001001.01.

Murphy, R. F. (2019). *Artificial intelligence applications to support K-12 teachers and teaching: A review of promising applications, opportunities, and challenges*. Rand Corporation. https://www.jstor.org/stable/resrep19907?seq=15.

Nautiyal, R., Albrecht, J. N., & Nautiyal, A. (2023). ChatGPT and tourism academia. *Annals of Tourism Research*, *99*. https://doi.org/10.1016/j.annals.2023.103544.

Noe, R. A. (2009). *Employee training and development*. McGraw Hill.

Noster, N., Gerber, S., & Siller, H. S. (2024). *Pre-Service Teachers 'Approaches in Solving Mathematics Tasks with ChatGPT–A Qualitative Analysis of the Current Status Quo*. https://doi.org/10.21203/rs.3.rs-4182920/v1.

Ocaña-Fernández, Y., Valenzuela-Fernández, L. A., & Garro-Aburto, L. L. (2019). Artificial intelligence and its implications in higher education. *Journal of Educational Psychology-Propositos y Representaciones*, 7(2), 553–568. https://doi.org/10.20511/pyr2019.v7n2.274.

OpenAI. (2022). *Introducing ChatGPT*. https://openai.com/blog/chatgpt/.

OpenAI. (2024). *Prompt engineering*. https://platform.openai.com/docs/guides/prompt-engineering.

Ouyang, F., Zheng, L., & Jiao, P. (2022). Artificial intelligence in online higher education: A systematic review of empirical research from 2011 to 2020. *Education and Information Technologies*, *27*, 7893–7925. https://doi.org/10.1007/s10639-022-10925-9.

Owens, B. (2023). How Nature readers are using ChatGPT. *Nature*, *615*(7950), 20. https://doi.org/10.1038/d41586-023-00500-8.

Pavlova, N. H. (2024). Flipped dialogic learning method with ChatGPT: A case study. *International Electronic Journal of Mathematics Education*, *19*(1), em0764. https://doi.org/10.29333/iejme/14025.

Pepin, B., Buchholtz, N., & Salinas-Hernández, U. (2025). A scoping survey of ChatGPT in mathematics education. *Digital Experiences in Mathematics Education*, *11*, 9–41. https://doi.org/10.1007/s40751-025-00172-1.

Popenici, S. A., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning*, *12*(22). https://doi.org/10.1186/s41039-017-0062-8.

Qadir, J. (2022). *Engineering education in the era of ChatGPT: Promise and pitfalls of generative AI for education*. https://doi.org/10.36227/techrxiv.21789434.v1.

Qi, X., Zhu, Z., & Wu, B. (2023). The promise and peril of ChatGPT in geriatric nursing education: What we know and do not know. *Aging and Health Research*, *3*(2). https://doi.org/10.1016/j.ahr.2023.100136.

Rawas, S. (2023). ChatGPT: Empowering lifelong learning in the digital age of higher education. *Education and Information Technologies*, 1–14. https://doi.org/10.1007/s10639-023-12114-8.

Rospigliosi, P. A. (2023). Artificial intelligence in teaching and learning: What questions should we ask of ChatGPT? *Interactive Learning Environments*, *31*(1), 1–3. https://doi.org/10.1080/10494820.2023.2180191.

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

Shukla Shubhendu, S., & Vijay, J. (2013). Applicability of artificial intelligence in different fields of life. *International Journal of Scientific Engineering and Research*, *1*(1), 28–35.

Siegerink, B., Pet, L. A., Rosendaal, F. R., & Schoones, J. W. (2023). ChatGPT as an author of academic papers is wrong and highlights the concepts of accountability and contributorship. *Nurse Education in Practice*, *68*. https://doi.org/10.1016/j.nepr.2023.103599.

Southgate, E., Blackmore, K., Pieschl, S., Grimes, S., McGuire, J., & Smithers, K. (2019). *Artificial Intelligence and Emerging Technologies in Schools*. University of Newcastle. https://researchoutput.csu.edu.au/en/publications/artificial-intelligence-and-emerging.

Taani, O., & Alabidi, S. (2024). ChatGPT in education: Benefits and challenges of ChatGPT for mathematics and science teaching practices. *International Journal of Mathematical Education in Science and Technology*, 1–30. https://doi.org/10.1080/0020739X.2024.2357341.

Tapan Broutin, M. S. (2023). Examination of Questions Asked by Pre-Service Mathematics Teachers in their Initial Experiences with ChatGPT. *Journal of Uludag University Faculty of Education*, *36*(2), 1–26. https://doi.org/10.19171/uefad.1299680.

Thurzo, A., Strunga, M., Urban, R., Surovková, J., & Afrashtehfar, K. I. (2023). Impact of artificial intelligence on dental education: A review and guide for curriculum update. *Education Sciences*, *13*(2), 150. https://doi.org/10.3390/educsci13020150.

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, *10*(1), 15. https://doi.org/10.1186/s40561-023-00237-x.

Wang, P. (2019). On defining artificial intelligence. *Journal of Artificial General Intelligence*, *11*(2). https://doi.org/10.2478/jagi-2019-0002.

Wardat, Y., Tashtoush, M. A., AlAli, R., & Jarrah, A. M. (2023). ChatGPT: A revolutionary tool for teaching and learning mathematics. *Eurasia Journal of Mathematics, Science and Technology Education*, *19*(7). https://doi.org/10.29333/ejmste/13272.

Wiliam, D., & Thompson, M. (2017). Integrating assessment with learning: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment shaping teaching and learning* (pp. 53–82). Routledge.

Xu, L. D., Lu, Y., & Li, L. (2021). Embedding blockchain technology into IoT for security: A survey. *IEEE Internet of Things Journal*, *8*(13), 10452–10473. https://doi.org/10.1109/JIOT.2021.3060508.

Yifan, W., Mengmeng, Y., & Omar, M. K. (2023). "a friend or a foe" determining factors contributed to the use of ChatGPT among University Students. *International Journal of Academic Research in Progressive Education and Development*, *12*(2), 2184–2201. https://doi.org/10.6007/IJARPED/v12-i2/17400.

Yin, R. K. (2009). *Case study research: Design and methods* (4th ed.). Sage.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–where are the educators? *International Journal of Educational Technology in Higher Education*, *16*(1), 1–27. https://doi.org/10.1186/s41239-019-0171-0.

Zhai, X. (2022). *ChatGPT user experience: Implications for education*. http://dx.doi.org/10.2139/ssrn.4312418.

Zhou, K., Yang, J., Loy, C. C., & Liu, Z. (2022). Learning to prompt for vision-language models. *International Journal of Computer Vision*, *130*(9), 2337–2348. https://doi.org/10.1109/cvpr52688.2022.01631.

Zhu, C., Sun, M., Luo, J., Li, T., & Wang, M. (2023). How to harness the potential of ChatGPT in education? *Knowledge Management & E-Learning*, *15*(2), 133–152. https://doi.org/10.34105/j.kmel.2023.15.008.