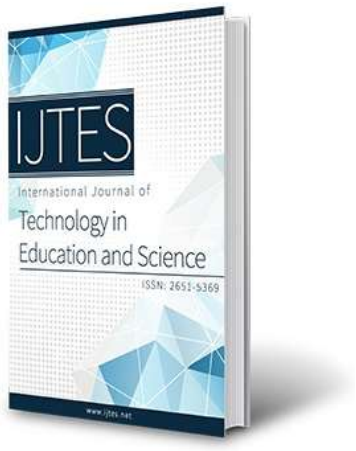

www.ijtes.net

Exploring Pre-Service Teachers' Engagement with Generative AI for Multiple-Choice Question Generation

Kübra Karakaya Özyer
Eskişehir Osmangazi University, Turkiye

To cite this article:

Karakaya Ozyer, K. (2025). Exploring pre-service teachers' engagement with generative AI for multiple-choice question generation. *International Journal of Technology in Education and Science (IJTES), 9*(3), 322-353. https://doi.org/10.46328/ijtes.624

IJTES International Journal of Technology in Education and Science
2025, Vol. 9, No. 3, 322-353 https://doi.org/10.46328/ijtes.624

Exploring Pre-Service Teachers' Engagement with Generative AI for Multiple-Choice Question Generation

Kübra Karakaya Özyer

Article Info

Article History

Received:
10 February 2025
Accepted:
20 June 2025

Keywords

Pre-service teachers
Generative artificial intelligence
AI in teaching
AI in education

Abstract

This study aims to delve into the process and perceptions of pre-service teachers as they engage in generating multiple-choice questions with the assistance of generative AI tools. Adopting a single-case study design, the research involved the participation of 35 pre-service teachers. The participants were tasked with utilizing generative AI tools to create multiple-choice questions and evaluate the generated items. Following this, a four-hour training session on item writing was provided to the participants. Post-training, they were asked to reassess the AI-generated products. The final phase involved collecting their reflective thoughts on this experience in writing. The qualitative data gathered were subjected to thematic analysis. A comparative examination of the participants' evaluations indicated a shift towards more detailed assessments that adhered more closely to measurement and evaluation standards. Reflective insights highlighted their positive perceptions of generative AI.

Introduction

In the 1960s, the character of Rosie, the robot maid from "The Jetsons" cartoon, represented a utopian vision of technological assistance, characterized by her helpfulness and emotional responses. This depiction, once considered purely speculative, aligns remarkably with the technological strides observed in the 2020s, especially in the realm of artificial intelligence (AI). The current era has witnessed AI becoming increasingly pivotal in numerous sectors, thereby realizing a level of human-like interaction that was previously envisioned as futuristic (Keleş, 2022). The application of AI technology has led to significant breakthroughs, from facilitating early disease diagnosis in the healthcare sector (Shaheen, 2021) to personalizing educational experiences, streamlining manufacturing processes (Shaikh et al., 2021), and enhancing financial decision-making (Cao, 2022; Ribeiro et al., 2021). These advancements have seamlessly integrated AI into the fabric of daily life, often beyond our immediate awareness. Particularly in education, AI has emerged as a fundamental element, highlighting its critical role and pervasive influence in the evolution of pedagogical and learning methodologies.

There are various generative AI applications which stands out for its potential to revolutionize the way assessments are designed, administered, and evaluated (Bozkurt et al., 2023; Łodzikowski et al., 2023; Zhang & Aslan, 2021). This technology, defined by its capacity to generate new content by identifying patterns within large datasets, presents unparalleled opportunities to improve the quality and efficiency of educational assessments.

The advent of generative AI tools such as Generative Pre-trained Transformers (GPT) has introduced novel methodologies for question generation, marking a significant departure from traditional assessment practices (Bhowmick et al., 2023; Mello et al., 2023).

This study wants to investigate the integration of generative AI and the assessment process. With the increasing reliance on technology in educational settings, the role of generative AI in crafting assessments that are both dynamic and reflective of individual learning paths is of paramount importance. By examining the experiences of pre-service teachers using generative AI for item generation, this study seeks to explore the complexities of integrating this technology, highlighting both the opportunities it offers and the challenges it poses. Through a comprehensive analysis, this study seeks to contribute to the broader discourse on the future of education, highlighting the transformative potential of generative AI in enhancing pedagogical practices and learning experiences.

Literature Review

Historical Development and Definition of Generative AI

Artificial Intelligence (AI) has evolved substantially since its mid-20th century origins, progressing from Turing's (1950) theoretical foundations to McCarthy et al.'s (1955) formal establishment of the field. This evolution has seen AI transition from abstract concepts to practical applications, culminating in advanced systems like Generative AI (GenAI). These sophisticated tools now perform tasks once exclusive to human cognition, marking a significant shift in computational capabilities and their potential impact on various domains.

Generative Artificial Intelligence (AI) is distinguished by its capacity to produce novel content by identifying and applying patterns from existing datasets, including text, images, audio, and video (Dwivedi et al., 2023; Kasneci et al., 2023; Lim et al., 2023). This technology transcends the boundaries of simple automated responses, undertaking a range of sophisticated tasks that demonstrate its versatility and complexity (Ding et al., 2023; Tlili et al., 2023). This capability has dramatically shifted how we interact with digital content, offering possibilities previously unimagined.

Generative Pre-trained Transformer architectures have demonstrated exceptional proficiency in generating contextually sophisticated textual compositions across diverse linguistic domains and scholarly genres (Aydin & Karaarslan, 2023; Ding et al., 2023). Furthermore, these models have successfully simulated realistic dialogs in diverse applications, from customer service chatbots to interactive characters in video games, showcasing their advanced language processing capabilities (Aydin & Karaarslan, 2023; Jovanović, 2023; Korngiebel & Mooney, 2021; Pavlik, 2023).

The launch of OpenAI's Chat GPT-3.5 model in November 2022 marked a significant milestone in the field of GPTs, achieving over one million users within the first five days of its release (Dans, 2023). Following the introduction of Chat GPT-3.5, other significant models such as Bing's Copilot and Google's Bard were introduced in 2023, further enriching the landscape of AI-driven tools available to users. This rapid adoption underscored the

global impact of generative AI technologies and highlighted their potential to revolutionize various sectors, including education.

GPT in Education

The integration of GPT technologies represents a paradigm shift in pedagogical approaches, substantially augmenting educational practices through advanced content generation, linguistic mediation, and comprehensive textual synthesis (Berg, 2023; Chan & Zhou, 2023; Tate et al., 2023; Lim et al., 2023). This technology enables the development of personalized learning materials and supports multilingual education, making complex information more accessible to students (Tzirides et al., 2023; Zastudil et al., 2023). AI-driven chatbots and platforms like ChatGPT, Google Bard, and Bing have demonstrated the capacity to tailor learning experiences, provide constructive feedback, and enhance student engagement (Cunningham-Nelson et al., 2019; Kadaruddin, 2023; Kuhail et al., 2023; Qadir, 2023; Su & Yang, 2023; Zawacki-Richter et al., 2019). Research indicates that AI-supported applications can significantly improve academic achievement and positively affect motivation in learning (Khan et al., 2021; Kim et al., 2021). Specifically, ChatGPT's ability to provide customized recommendations and serve as a virtual tutor underscores its capacity for delivering targeted instruction tailored to the unique requirements of each student (Su & Yang, 2023). This personalized approach aligns with broader findings on AI in education, as emerging evidence suggests that such tailored interventions can have significant impacts on student outcomes. Research shows that AI-supported applications significantly support academic achievement and positively affect motivation and engagement in learning (Khan et al., 2021; Kim et al., 2021; Zheng et al., 2023; Xia et al., 2022). While the benefits for students are evident, the potential of GPT and AI-driven educational tools extends beyond enhancing student learning, offering significant advantages for educators in streamlining their teaching processes and administrative tasks.

ChatGPT facilitates teachers in efficiently managing student inquiries, thereby optimizing the allocation of time to essential educational responsibilities (Lin, 2022; Durall & Kapros, 2020; Su & Yang, 2023). Tools such as Gradescope and Autolab enhance the monitoring of student progress, enriching the learning experience by advancing educational models, assessment systems, and the overall educational ecosystem (Baker & Smith, 2019; Okonkwo & Ade-Ibijola, 2020; Su & Yang, 2023). The capability of generative AI to support educators in tasks such as essay writing and providing thematic guidance is particularly noteworthy, offering significant insights to administrators by facilitating closer monitoring of students' strengths and weaknesses (Chan & Zary, 2019; Karsenti, 2019; Villegas-Ch et al., 2021). Concurrently, the trajectory of artificial intelligence applications within educational domains is progressively advancing toward increasingly complex and nuanced implementation strategies. Arizona State University's initiative to design a virtual faculty member using the GPT-4 language model for classroom instruction aims to further personalize the learning experience for students, indicating a forward leap in educational methodologies (Gizchina.com, 2024).

While the emergent integration of generative artificial intelligence technologies within educational contexts presents potentially transformative pedagogical methodologies, it is critically essential to conduct rigorous, comprehensive evaluations that systematically examine technological efficacy, validate data reliability, and

comprehensively scrutinize the multifaceted ethical dimensions inherent in such technological interventions (Kadaruddin, 2023; Qadir, 2023). The cumulative research underscores the potential benefits of AI in educational contexts, but cautions against the need for a judicious evaluation of its ethical and practical ramifications.

Generative artificial intelligence represents a pivotal methodological innovation in educational assessment and evaluative paradigms (Crompton & Burke, 2023). Generative AI's integration into educational assessments is enabling the creation of more sophisticated, varied methods of question generation and evaluation that closely align with educational outcomes. Specifically, the development of AI-powered tools such as Learnosity, Gradescope, and Education Copilot has marked a paradigm shift in how educators prepare and assess examinations. For instance, Learnosity and Education Copilot, are equipped with capabilities to generate items (questions) (e.g. Multiple-choice, Fill in the blanks etc.) directly tied to specific learning objectives (e.g. Mathematics, Chemistry), facilitating a targeted evaluation of student learning. Generative AI is utilized not only in question generation but also in the evaluation of exams. For example, Mizumoto and Eguchi (2023) conducted a study which underscores the potential of ChatGPT as a reliable and valid tool for the automated scoring of essays. Gradescope.com and Assess.com also offer an innovative solution for the automatic grading of student responses, significantly enhancing the efficiency of the evaluation process. These platforms allow multiple-choice, fill-in-the-blank or essay-type questions to be evaluated with the support of AI. These technological interventions support educators by streamlining the exam preparation and grading process, thereby allowing for a more effective allocation of instructional time (Kaplan-Rakowski et al., 2023).

While generative AI offers promising applications in educational contexts, its implementation is not without significant challenges and potential drawbacks that warrant careful consideration. One of the potential drawbacks is academic integrity. The ability of GPT models to produce human-like text has introduced new challenges in detecting plagiarism within student assignments (Cotton et al., 2023). The nearly flawless capability of these AI tools to generate text, images, videos, and presentations has led some educational institutions, such as New York City and Los Angeles Unified Schools, to restrict internet access to ChatGPT (Korn & Kelly, 2023; Shen-Berro, 2023). However, every day brings a new set of generative AI tools, and it becomes harder to detect their use. Consequently, some institutions have felt the need to publish guidelines for ChatGPT's usage (The McGraw Center for Teaching & Learning, Princeton University, 2023). AI-generated content faces challenges in terms of accuracy and reliability. The lack of transparency in AI's data sources and processing methods can sometimes lead to erroneous outputs (Islam & Islam, 2023; Kasneci et al., 2023). Furthermore, studies have identified a tendency for AI to produce biased responses, as demonstrated by Bolukbasi et al. (2016), highlighting the critical need for oversight in its educational application.

Generative AI's role in educational assessment is multifaceted, offering the dual benefits of operational efficiency and pedagogical depth. By automating routine tasks and enabling more nuanced assessments, these technologies foster a more personalized and responsive educational environment. Considering the positive and negative effects of generative artificial intelligence in assessment and evaluation processes in education, the continued exploration of generative AI in the assessment process promises to further enhance the quality and effectiveness of teaching and learning experiences.

Educator Competencies and Perspectives on GPT

The Technological Pedagogical and Content Knowledge (TPACK) framework provides a crucial roadmap for effectively integrating emerging technologies into education. This approach emphasizes the importance of teachers blending their knowledge of technology, pedagogy, and subject matter to enhance learning objectives and curriculum integration (Koehler & Mishra, 2009). Research supports the effectiveness of this framework, showing that its application can lead to improved academic achievement and problem-solving skills among students (Altun & Usta, 2019; Wetzel & Marshall, 2011).

Central to the success of TPACK is the technological proficiency of teachers. A study by John (2005) demonstrated that as educators improve their tech skills, they become more adept at incorporating digital tools into their classroom activities. This increased proficiency is particularly important when considering advanced technologies like GPT in education. GPT exemplifies the potential of TPACK in action, offering opportunities for personalized and innovative learning experiences. For example, a math teacher could utilize ChatGPT to design engaging lessons on complex topics such as functions. By aligning this technology with specific pedagogical goals and subject content, teachers can create a more dynamic learning environment that students feel engagement.

The successful implementation of GPT and similar technologies in education, therefore, depends on more than just the availability of these tools. It requires teachers to strategically incorporate them into their teaching practices, in line with TPACK principles (Koehler & Mishra, 2009). This approach ensures that technology serves as a powerful enhancer of education rather than a mere add-on, potentially revolutionizing how we approach teaching and learning in the digital age.

In addition to teachers' proper integration of technology into their lessons, their attitudes toward these technologies are also important. Teachers who are open-minded and enthusiastic about integrating technology into their lessons often create a more engaging and dynamic learning environment for their students (Christensen, 2002; Lambert & Lane, 2004; Liu, 2011). Their positive attitudes toward technology can inspire students to become more curious and comfortable with using digital tools for their learning (Arkhipova et al., 2019). Furthermore, according to the Technology Acceptance Model, teachers' attitudes toward technologies affect their intentions to use those technologies in the future (Nair & Das, 2012). This situation is also valid for pre-service teachers. Studies have revealed that pre-service teachers' intention to use technologies in the future is directly related to their attitudes towards those technologies (Teo, 2012; Teo & Zhou, 2017). Both teachers' and pre-service teachers' perceptions of ease of use and usefulness of technology tend to lead them to use certain technologies more in their professional lives (Teo, 2011). Joo et al. (2018) explored the connections between pre-service teachers' TPACK, self-efficacy, perceived ease of use, and usefulness of technology, revealing that TPACK significantly impacts self-efficacy and perceived ease and usefulness of technology in education. In addition, the study found that preservice teachers' self-efficacy, perceived ease of use, and usefulness of technology directly influence their intentions to use technology. The cultivation of TPACK among pre-service educators demonstrates a significant correlation with enhanced technological self-efficacy and pedagogical integration strategies. Consequently, it is imperative for

educational stakeholders—including teacher preparation institutions, policy formulators, and academic researchers—to systematically prioritize comprehensive professional development initiatives that substantively enhance educators' technological competencies, attitudinal orientations, and instructional integration capacities (Agyei & Voogt, 2011; Ertmer et al., 2012).

Kuleto et al. (2022) executed an empirical study among K-12 educational practitioners in Serbia, focusing on comprehensively mapping their existing artificial intelligence knowledge, current proactive technological integration strategies, and critically examining the divergences between their perceptual landscape and technological expectations. The findings highlighted a positive link between teachers' experiences with AI in Education (AIED) and their views on it, which subsequently affected their intentions to use AI in teaching. This indicates that enhanced professional development enabling hands-on AI use could increase teachers' readiness to integrate AI into their routine teaching practices. Given the evolving nature of generative AI such as ChatGPT, this topic becomes pertinent for current and future educators. In order to reveal this situation, Yang and Chen (2023) explored how pre-service teachers' beliefs and training experiences shape their eventual teaching methods, focusing on 26 pre-service information technology educators in Taiwan. Through questionnaires, interviews, and chatbot interaction logs, it was found that these educators leveraged chatbots for various learning activities, including organizing knowledge, content discovery, and non-curricular discussions. Their study claims that such interactions with chatbots might influence educators' decisions on incorporating or omitting certain activities in a Learning Management System (LMS).

The integration of generative AI into education has sparked diverse reactions among educators, ranging from enthusiasm about its transformative potential to concerns regarding academic integrity and the changing role of teachers (Chounta et al., 2022; Choi et al., 2023). This varied landscape underscores the need for a deeper understanding of how these technologies can be effectively harnessed for pedagogical innovation.

Significance of the Current Study

The significance of this study is multifaceted, addressing critical gaps in the literature and practice of integrating generative AI in pre-service teacher training. While existing research has predominantly focused on the technical capabilities and potential efficiencies of AI tools (Kaplan-Rakowski et al., 2023; Mizumoto and Eguchi, 2023), there remains a notable lacuna in understanding their practical integration into teaching and learning processes, especially within pre-service education contexts. This research makes several key contributions:

1. Practical Application in Pre-service Training: By exploring the use of generative AI in developing multiple-choice questions, this study provides much-needed empirical evidence on how AI can enhance assessment skills among pre-service teachers. This practical approach offers valuable insights into the challenges and benefits of AI integration in teacher education, paving the way for more effective AI-integrated pedagogical strategies.
2. Bridging Theory and Practice: The focus on experiential learning with generative AI, particularly in creating educational materials and assessment questions, addresses the crucial need to connect theoretical knowledge with practical application. This bridge is essential for preparing future educators to effectively utilize AI in real-world teaching scenarios.

3. Addressing Technological Proficiency: As generative AI becomes increasingly central to professional development and modern education (Baidoo-Anu & Ansah, 2023; Sengupta & Chakraborty, 2020), this study directly addresses the growing need for future educators to be proficient in advanced technologies. It provides a framework for how pre-service teachers can gain hands-on experience and comprehensive understanding of these tools (Lo, 2023).
4. Enhancing Assessment Practices: By focusing on the development of multiple-choice questions, the study contributes to the broader field of educational assessment. It explores how AI can potentially revolutionize the way teachers approach assessment design, potentially leading to more efficient and effective evaluation methods.
5. Informing Policy and Curriculum Development: The findings from this study have the potential to inform policy decisions and curriculum development in teacher education programs. By providing empirical evidence on the benefits and challenges of AI integration, it can guide institutions in updating their pre-service training curricula to include AI literacy and practical skills.
6. Ethical Considerations in AI Use: This research also touches upon the ethical implications of using AI in education, contributing to the ongoing discourse about responsible AI use in teaching and assessment. It provides a foundation for developing guidelines on ethical AI integration in educational practices.
7. Long-term Impact on Educational Quality: By enhancing pre-service teacher training programs, this study ultimately aims to improve the quality of education in future classrooms. Better-prepared teachers, equipped with AI skills, can potentially lead to enhanced educational delivery and increased student engagement.

In conclusion, this study not only contributes to the academic discourse on AI in education but also has far-reaching practical implications. It addresses the critical need for empirical research in this rapidly evolving field, offering insights that can shape the future of teacher education and, by extension, the broader landscape of educational practices in an AI-integrated world.

Research Objectives

The main objective of this study is to examine pre-service teachers' question generation experiences with generative AI in detail. In this context, we aim to understand and evaluate these experiences by focusing on the following research questions:
1. How did pre-service teachers evaluate the multiple-choice item types they produced using generative AI before they were trained on the subject?
2. How did the pre-service teachers evaluate the multiple-choice item types they produced using generative AI after the training on the subject?
3. How do pre-service teachers evaluate their item generation experiences using generative AI?

Method

Research Design

In this study, a single case study, a qualitative research method, will be preferred to examine pre-service teachers' experiences of generating questions using generative artificial intelligence (AI). This method allows a researcher

to describe in detail and in depth a limited system, i.e. a situation from real life (Creswell, 2012). The case in the current study is the process of pre-service teachers' generating multiple-choice questions using generative AI and their evaluation of these generated questions.

Participants

This research employed a purposive sampling method to select participants who could significantly contribute to the findings. After the ethics committee report of the research was obtained, pre-service teachers, enrolled in the Faculty of Education at a state university located in the Central Anatolia region of Türkiye, were recruited for the study through direct communication and announcements made within the faculty. Since the assessment and evaluation course is given in the 3rd grade in this faculty, the participants were selected from this group. Those who wanted to contribute to the research were selected by distributing the voluntary participation form. Although the research started with 37 individuals, it was noted that two participants did not complete their second evaluations. Consequently, these individuals were excluded from the analysis, and 35 people are left. Two of the participants were men and 33 were women.

Procedure

This study is designed in four phases in the fall semester of 2023-2024 (see Figure 1). In the first stage, the participants were asked to choose an education level (primary, secondary), a subject, and a unit. The participants were asked to prepare one multiple-choice question using the help of generative AI in accordance with the learning outcomes of the unit they selected. To standardize the training each participant received, the study implemented a fixed training schedule. The participants were required to fill in the relevant sections of the document (see Appendix 2) within a 2-week period. At the end of 2 weeks, these documents were collected via online platform.

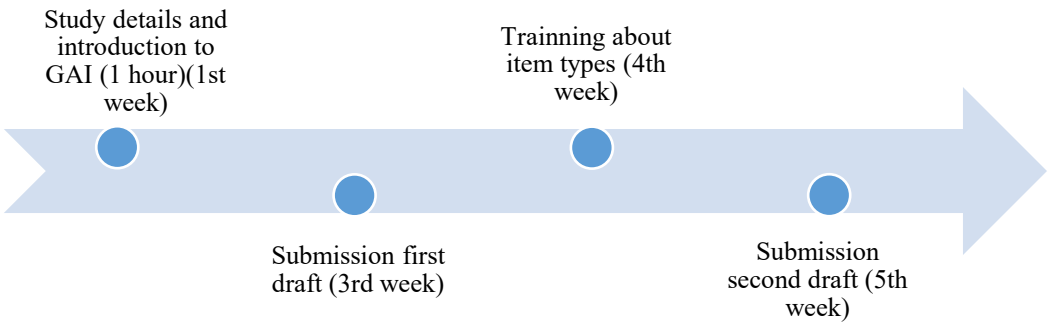

Figure 1. Research Process

The chosen class level, subject, and artificial intelligence tool used by the participants are detailed in Table 1. The table lists 35 participants (P1 to P35) who use generative AI tools, specifically ChatGPT and Google Bard, across various school grades (ranging from 3rd to 12th) and subjects including Turkish language, Science, English language, History, History of Turkish Revolution (İnkılap), Social Science, Geography, religious culture and ethics, Life Science, Turkish Literature, and Philosophy. ChatGPT appears to be the most frequently used tool

across a wider range of grades and subjects. It is notably prevalent in Turkish, English, History, and Geography classes.

Table 1. Details of the Participants' Preferences for the Study

Variables	Categories	Participant codes
Grade level	3th	P2, P22, P32
	4th	P4, P8, P13, P33
	5th	P3, P15, P23, P29, P34
	6th	P18
	7th	P1, P9, P12
	8th	P6, P7, P10, P16, P20
	9th	P5, P14, P24, P28, P35
	10th	P11, P19, P25, P26, P27, P30
	11th	P17, P21
	12th	P31
Subject	Turkish Language	P1, P3, P6, P10, P12, P13, P15, P16, P17, P21, P23, P33, P34
	Geography	P11, P27, P28, P30, P35
	History	P5, P14, P19, P26
	English Language	P4, P18, P24,
	Science	P2, P20, P32
	Social Science	P9, P29
	History of Turkish Revolution	P7, P31
	Religious Culture and Ethics	P8
	Life Science	P22
	Philosophy	P25
Generative AI Tool	ChatGPT	P1, P2, P4, P5, P6, P7, P9, P10, P12, P13, P14, P15, P16, P17, P18, P19, P20, P23, P24, P25, P26, P27, P28, P29, P30, P32, P33, P34, P35
	Google Bard	P3, P8, P11, P21, P22, P31

To ensure that the participants' responses and interactions with generative AI remained uninfluenced and natural, the study did not provide detailed training on the technology. Instead, we briefly introduced what generative AI is and how it works and explained the expectations of the participants. This approach will allow participants to freely use their own strategies during the study process and gain first-hand experience with generative AI. In this way, the research results will provide a more realistic perspective on students' ability to learn on their own and their interaction with generative AI.

After the participants made their first submissions, they received 4 hours of training on item types. In this training, the definitions and characteristics of item types within the scope of educational assessment and evaluation, the situations to be considered when writing items, and the relationships of items with validity, reliability, and usefulness were specified. After these trainings, the participants were asked to re-evaluate the questions they had created using artificial intelligence. Participants recorded their re-evaluations and submitted them online.

Data Collection Tools

In the initial stage of the study, the participants were provided with instructions (outlined in Appendix 2) to create multiple-choice items using generative AI tools such as ChatGPT, Google Bard, and Bing, etc. The instruction document required participants to record the specific prompts used to generate questions and to include transcripts of interactions with AI chatbots. This process elucidated the strategies employed by pre-service teachers in generating questions with generative AI and to gain insights into their hands-on experiences using these technologies.

Upon completion of item generation, the participants independently evaluated the quality and relevance of the generated items without any predefined guidance. Following this evaluative process, participants were prompted to reflect on and document their overall experiences and perceptions related to using generative AI for educational purposes. This reflective exercise was intended to capture participants' subjective evaluations and emotional responses to their interaction with generative AI technologies throughout the study.

Data Analysis

Thematic analysis was employed as the primary methodological approach. Thematic analysis, as outlined by Braun and Clarke (2006), serves to dissect the content of texts or documents, facilitating the extraction of significant themes by identifying recurrent patterns, expressions, or subjects within the data. This analytical process is pivotal for distilling the essence of textual content and unveiling the underlying ideas.

Creswell (2012) outlines the qualitative research process as organizing and preparing data, conducting initial explorations and coding to distill descriptions and themes, visually representing and narratively reporting findings, interpreting their significance in relation to existing literature and personal insights, and validating the findings through accuracy checks such as member checking, triangulation, and auditing. To conduct the thematic analysis of qualitative data within this study, the researcher engaged in multiple readings of the dataset to gain a thorough comprehension. The subsequent step involved delineating the analytical units, followed by the generation of codes. These codes were developed by scrutinizing the prompts used by participants for question generation, with a focus on the nuances of prompt engineering.

The thematic analysis was performed distinctly to the initial and final evaluations made by participants regarding the generated items. The creation and refinement of codes for the initial evaluations benefited from expert consultation, ensuring their validity and relevance. These refined codes were then meticulously reviewed to

establish coherent themes. A similar procedural approach was adopted for analyzing the participants' final evaluation responses, thereby ensuring a robust and methodologically sound analysis of the qualitative data collected throughout the study.

Validity and Reliability

Assessment and educational technology experts rigorously evaluated the data collection tools developed for this study to ensure their validity and effectiveness. Expert feedback was meticulously incorporated to refine the tools to their final form. Additionally, a comprehensive description of the research process was provided to enhance the credibility of the study, with efforts made to meticulously detail each step involved. The credibility was further bolstered by employing triangulation, using a combination of data collection methods (item evaluations, reflective thoughts, and screenshots) to enrich the dataset and reinforce the findings.

In addition, an assessment expert was consulted for the analysis of participant evaluations regarding item types. The emerging codes and themes were discussed with this expert, and consensus was reached on the codes and themes related to the assessment expertise. This collaborative and iterative process not only enhanced the internal consistency of the coding but also ensured that the study's findings were grounded in expert validation, thereby strengthening the overall credibility and validity of the research.

Ethical Aspects

Ethical standards were taken into account in this study. Prior to data collection, ethical approval was obtained from the Eskişehir Osmangazi University's Ethics Review Board (Approval ID: E-64075176-050.04-2400023969). Before data collection, participants were given a voluntary participation form and data were collected from those who gave consent. Data were encrypted and stored securely, with participant anonymity preserved through unique code names (e.g., P1, P2, ... P35), reflecting the study's dedication to ethical integrity and privacy.

Results

Thematic Analysis of Pre-Service Teachers' Evaluations of AI-Generated Multiple-Choice Questions

In this study, pre-service teachers were asked to evaluate the multiple-choice questions (MCQ) they created using artificial intelligence. The participants were evaluated in both the second and the last phase of the study. The initial evaluations made in the second phase of the study were subjected to thematic analysis. The thematic analysis of participant evaluations revealed significant insights into the effectiveness of AI-generated questions and, their pedagogical suitability (see Table 2). Table 2 presents the initial evaluations of AI-generated questions by the participants. The table uses specific coding to indicate varied responses to the effectiveness and suitability of the questions:

- '+' signifies a positive assessment or an aspect that was favorably received.
- '-' denotes a negative assessment or an aspect that was unfavorably received.

Table 2. Results of Thematic Analysis of Pre-Service Teachers' Initial Evaluations

Themes	Codes	Participants
AI Effectiveness in Question Generation	Item Complexity	Basic: P3, P5, P6, P8, P13, P16, P21, P34, P35 Medium: P20, P23, P27, P28 Hard: P19, P29, P33
	Cognitive Level	P9(Knowledge)
	Content Relevance	P6, P7, P8, P9, P10, P14, P20, P21, P22, P23, P26, P27, P29, P30
	Suitability of Grade Levels	P2, P4, P6, P8, P10, P12, P14, P22, P32
	Instruction Responsiveness	P1(+), P5(-), P10(-), P11(-), P16(-), P17(-), P21(+), P24(-), P26(-), P27(+), P28(+), P29(+), P31(+), P32(+), P33(-)
Pedagogical Suitability of AI-Generated Questions	Item Clarity	P2, P5, P6, P7, P14, P23, P25, P30
	Comprehensive Coverage	P1, P20
	Discriminative Power of Items	P12, P33, P34
	Suitability of the Options	P12, P15(-), P19(-), P21(-), P22(-), P25(-), P26(-), P31(+), P33(-)
	Item Format Suitability	P4, P21

According to Table 2, the thematic analysis revealed two themes: 'AI Effectiveness in Question Generation' and 'Pedagogical Suitability of AI-Generated Questions'. Under the theme 'AI Effectiveness in Question Generation', the codes 'Item Complexity', 'Cognitive Level', 'Content Relevance', 'Suitability for Grade Levels', and 'Instructional Responsiveness' have been identified.

In evaluating the MCQs, the participants' feedback revealed varied perspectives on the complexity, cognitive level, content relevance, suitability of grade level, and instructional responsiveness of the AI-generated items. The complexity of the questions spanned from basic to hard, with a notable number of participants (P3, P5, P6, P8, P13, P16, P21, P34, P35) finding the questions to be of a basic level and appreciating their clarity and directness. This suggests that while AI-generated questions are accessible to various learners, there is a trend toward generating items that test lower-order cognitive skills, as evidenced by P13's observation of items primarily assessing basic knowledge. P13 claimed that "I got a good answer to the question I asked, but the questions seemed simple for 4th grade. He could have prepared more intermediate-level questions." When the thematic analysis results were examined, it was determined that only the P9 participant made an evaluation on cognitive level. The participant said that she wanted the productive AI to prepare questions appropriate to the level of knowledge in

Bloom's taxonomy and that the question formed was appropriate for this level.

Content relevance emerged as a strong suit of AI-generated questions, with a significant cohort of participants (P6, P7, P8, P9, P10, P14, P20, P21, P22, P23, P26, P27, P29, P30) acknowledging their alignment with the curriculum, enhancing the utility of AI in crafting syllabus-aligned assessment tools. Comments on the suitability of questions for specific educational levels further illustrated this alignment, indicating AI's capacity to tailor content to various classroom contexts. For example, Participant 8 showed the appropriateness of the item for the grade level with the statement "I found the question very understandable and logical. It was a very simple question that can be answered very easily at the 8th grade level.".

However, instructional responsiveness—a measure of how well questions align with prompts—elicited mixed feedback. Positive responses (P1, P21, P27, P28, P29, P31, P32) highlighted AI's potential to understand and process prompts effectively. Participant 1 stated that the generative AI acted in accordance with its purpose with the statement "I asked it to prepare a multiple-choice question on the topic of "Forms of Expression", which is one of the topics of the unit of Meaning in the Piece, and I think that the question that emerged with the command I gave was compatible.". In contrast, some participants (P5, P10, P11, P16, P17, P24, P26, P33) expressed concerns about the limitations of AI in providing inaccurate and incomplete information. This critique, exemplified by P10's feedback, points to a gap in AI's ability to generate questions that challenge students beyond surface-level understanding. Such feedback underscores the need for AI technologies to evolve in their capacity to comprehend and incorporate advanced cognitive skills into question generation, ensuring that AI-generated MCQs can adequately reflect and assess the full spectrum of educational objectives.

The second theme that emerged from the participants' initial evaluations was 'Pedagogical Suitability of AI-Generated Questions'. Under this theme, the codes 'Item Clarity', Comprehensive coverage, discriminative power of items, suitability of options, and Item Format Suitability were collected.

Participants acknowledged the clarity of AI-generated items as a critical factor in facilitating effective assessment, with several (P2, P5, P6, P7, P14, P23, P25, P30) noting its role in eliminating ambiguity and directly assessing student knowledge, as highlighted by P7. The ability of items to cover comprehensive content and their discriminative power were also discussed, showcasing AI's potential in creating valuable educational assessments. However, the responses to the suitability of answer options were mixed, with some participants expressing dissatisfaction (P15, P19, P21, P22, P25, P26, P33) regarding the plausibility and relevance of distractors, while others (P31) viewed them positively. This divergence underscores the challenges faced in crafting effective distractors and the necessity for careful review. Additionally, the item format's structure and presentation received approval from certain participants (P4, P21), indicating AI's capability to generate well-structured questions. However, the varied feedback also emphasizes the need for educator involvement in refining and contextualizing AI-generated questions to ensure their educational efficacy and alignment with learning objectives.

In the second phase of the study, after providing training on creating multiple-choice items to pre-service teachers,

a nuanced reassessment of AI-generated questions was conducted. This phase was instrumental in uncovering shifts in perceptions, highlighting both the potential and limitations of AI in educational settings, as informed by direct participant evaluations (see Table 3). Table 3 details the post-training evaluations of the AI-generated questions, where participants reassessed the items after receiving further instruction on item generation and evaluation:

- '+' denotes a strong positive reassessment, indicating significant improvement or satisfaction.
- '-' signifies a strong negative reassessment, indicating significant concerns or dissatisfaction that persisted or intensified post-training.
- '+, -' are representing mixed or evolving perceptions among the participants.

Table 3. Results of Thematic Analysis of Pre-Service Teachers' Final Evaluations

Themes	Codes	Participants
AI Effectiveness in Item Generation	Item Complexity	P2 (easy), P5 (easy), P8(easy), P30(medium),
	Cognitive Level (Bloom's Taxonomy)	P1(basic), P3, P10(basic), P15(basic), P33(complex),
	Instruction Responsiveness	P1(-), P11(-), P16 (+), P33 (-)
	Suitability to the Student Level	P25, P32
	Conformity with Educational Objectives	P30
Pedagogical Suitability of AI-Generated Items	Item Clarity	P1, P4, P5, P9, P26, P30, P31, P32
	Identifying the Type of the Item	P1, P2, P4, P6, P8, P9, P11, P13, P14, P17, P21, P26, P29, P31, P33
	Suitability of Options and Distractors	P1(+), P2(-), P3(+), P4(+), P8(+), P9(+), P10(+), P12(+), P13 (+), P14(-, +), P15(-), P16(-), P17(+), P18(-), P19(-), P21(-), P22(+), P25(+), P31(+, -), P32(+), P33(-), P34(+), P35(-)
	Appropriateness of the Item Stem	P1(-), P2(+, -), P3(+), P6(+), P7(+), P9(-), P10(+), P11(+), P12(+), P13(+), P14(+), P15(+, -), P17(+, -), P20(-), P21(+), P22(+), P23(+, -), P24(+, -), P27(-), P28(-), P29(+, -), P32(+), P33, P34(+), P35(-)
	Existence of Instruction	P8(-), P21, P22(-), P27(-), P35(-)
	Each Item is Directed Toward a Learning Outcome	P2(+), P6(+), P24(+)
	Independence of Items from Each Other	P2(-)
Key Features of the Assessment Tool	Content Validity	P2(+), P4(+), P10(+), P25(+), P30(+), P31(+), P34(+)
	Reliability	P10(+), P19(-), P25(+), P31(+)
	Usability	P1, P31

The second evaluation showed that the pre-service teachers engaged in more detailed evaluations of the questions. In the final evaluation, in addition to the themes of ‘AI Effectiveness in Item Generation’ and ‘Pedagogical Suitability of AI-Generated Items’, a new theme for 'Key Features of the Assessment Tool' was created. Although the themes in the second evaluation are similar to those in the first evaluation, they differ in content. When the codes collected under the themes were examined carefully, it was revealed that the participants had fewer statements regarding the theme "AI Effectiveness in Item Generation". When the codes in the theme "Pedagogical Suitability of AI-Generated Items" were examined, it was determined that more detailed and appropriate evaluations were made in accordance with the assessment and evaluation rules.

Under the theme of ‘AI Effectiveness in Item Generation’, there were codes of Instruction Responsiveness, Item complexity, Cognitive level (Bloom’s Taxonomy), Suitability to student level, and Conformity with Educational Objectives. At first, the complexity and cognitive level of the items generated by AI were areas of particular interest. Participants P2, P5, and P8 described the items as simple, a perspective that may reflect an increased ability to critically evaluate question difficulty post-training. Participant P30 noted a medium level of complexity, suggesting variability in the perceived challenge posed by AI-generated questions. The training also seemed to influence perceptions of cognitive engagement, with participants identifying questions that spanned from basic knowledge to advanced cognitive processes, indicating the breadth of potential that AI might offer in targeting Bloom's taxonomy.

Following the training, the participants exhibited a more critical stance toward the instructional responsiveness of AI-generated questions. For instance, Participant P1 expressed dissatisfaction, noting that all responses seemed to fit a generic narrative, lacking the instructional depth they had come to appreciate post-training: "I find it insufficient with my current knowledge that all of the answers fall under 'narration'." This sentiment was echoed by others, such as Participant P11, who also indicated a negative shift in perspective, underscoring the challenges in aligning AI-generated content with specific educational goals.

Despite these critiques, the training appeared to enhance the understanding and appreciation of certain aspects of AI-generated items. Some participants, such as P25 and P32, recognized that the generated questions were appropriate for the level of the students. Besides, only one participant expressed an opinion on the suitability of the multiple-choice item created by artificial intelligence to the learning outcomes of the unit. Because the appropriateness of the question to the outcome is one of the indispensable features of an item, it can be concluded that the awareness of pre-service teachers on this issue is low.

In the refined assessment of AI-generated questions’ pedagogical suitability following participant training, nuanced evaluations emerged, particularly in the realms of item clarity, item type determination, distractor appropriateness, stem writing suitability, instruction presence, item independence, and each item’s alignment with specific learning outcomes. These evaluations, grounded in participant feedback, offer a multifaceted view of the pedagogical utility of AI in item generation.

Item clarity stood out as a strength, with participants such as P1 and P30 noting the unambiguous nature of the

questions, which facilitates a straightforward assessment of student knowledge. P1 remarked on the clear and direct formulation of questions, highlighting their potential to engage students without confusion. This clarity is crucial in educational settings, where the goal is to assess understanding rather than navigational ability through ambiguously worded items.

In the literature, multiple-choice items are categorized into various types according to item stems or options. For example, if there is only one correct answer in the options, this type is called the multiple-choice item type with one correct answer. Because of the thematic analysis, it was determined that some of the participants (P1, P2, P4, P6, P8, P9, P11, P13, P14, P17, P21, P26, P29, P31, P33) made evaluations using this information. This observation underscores a significant development in participants' evaluative responses following the training session, reflecting an enhanced application of their pedagogical knowledge in assessing the items.

During the training session, pre-service teachers were instructed on key considerations for crafting options and distractors for multiple-choice questions (MCQs). Leveraging this guidance, they re-assessed the items, leading to varied feedback on the appropriateness of the options and distractors. This divergence underscores the complexity inherent in designing effective MCQs. Positive evaluations from participants such as P1 and P3 highlighted generative AI's capacity to produce relevant and challenging distractors, thereby augmenting the pedagogical merit of the questions. On the contrary, critique from others, such as P2 and P16, underscored moments when distractors proved overly simplistic or misaligned, casting shadows on the generative AI's proficiency in delivering precise and high-quality options. This mix of positive and negative feedback underscores the challenges and considerations involved in employing generative AI for educational content creation, particularly in the nuanced task of MCQ development.

Stem writing similarly garnered mixed reviews. Participants like P3 and P6 praised the AI for generating concise and relevant stems, while others, such as P1 and P9, identified inconsistencies in stem quality, suggesting areas for AI improvement. The ability to generate well-constructed stems is fundamental because, it sets the stage for the question, guides students' thought processes, and ensures alignment with the intended learning objectives.

The presence of instructions within the questions was another area of critique, with some participants noting a lack of clear guidance on how to approach the questions (P8, P22, P27, P35). This omission can lead to confusion and misinterpretation, underscoring the importance of a comprehensive question design that includes explicit instructions when necessary.

Other considerations when creating multiple-choice items are that the items should be answered independently of each other and each item should address a single learning outcome. Item independence was also evaluated. While P2, P6, and 24 acknowledged the AI's ability to generate questions that stand alone and directly relate to learning outcomes, pointing to a targeted approach in question generation, only P2 expressed concern that the items could not be answered independently of each other.

The essential attributes of an effective assessment tool include reliability, validity, and utility, which were

emphasized by pre-service teachers following their training on item construction. Seven participants (P2, P4, P10, P25, P30, P31, P34) positively remarked on the content validity of the items generated, with one highlighting, "There is content validity because questions encompass all aspects related to the subject," indicating that AI was successful in creating questions that upheld content validity. Additionally, three participants (P10, P25, P31) believed in the high reliability of the items due to their objective scoring potential. Regarding utility, opinions varied; P1 critiqued the low practicality of item preparation, whereas P31 lauded the ease of application, demonstrating diverse perspectives on the usability of AI-generated assessment items.

A comparison of the two evaluations highlights a shift in focus and depth of understanding among the participants. While the initial evaluation centered on general impressions and AI's efficacy in item generation, the post-training evaluation adopted a more technical and detailed approach. The training appears to have enhanced participants' abilities to critique item structure and the characteristics of assessment tools, leading to more informed and nuanced evaluations. This transition illustrates the impact of educational interventions on participants' competencies in item design and assessment, underlining the importance of targeted training in enhancing analytical skills in educational settings.

Thematic Analysis of Pre-Service Teachers' Reflections on Using Generative AI for MCQ Creation

This study explored the experiences of 35 pre-service teachers as they engaged with generative AI to create multiple-choice questions, followed by reflection on their experiences. The thematic analysis of their reflective thoughts revealed a complex interplay of previous experiences, perceived positive aspects of generative AI, concerns about using AI in education, and insights related to technological, pedagogical, and content knowledge (TPACK), as well as future orientations toward the use of generative AI in educational settings (see Table 4).

Table 4. Thematic Analyses of Pre-service Teachers' Reflections on Using Generative AI for MCQ Generation

Themes	Sub-themes	Codes	Participants
Previous Experiences		Perception Shift	P1, P7, P32
		AI Familiarity Comfort	P5, P6, P16
		AI Unfamiliarity Barrier	P4, P19
Positive Aspects of Generative AI	Positive Features of The Technology Itself	Perceived Ease of Use	P6, P8, P20, P21, P29, P30, P31, P33
		Dialogic Interface	P15, P22, P25
		Option Diversity Perk	P1, P12, P21
		Language Adaptability	P5
	Positive Features of the Research Experience	Perceived Usefulness	P2, P3, P9, P10, P11, P13, P14, P15, P19, P20, P22, P23, P24, P28, P30, P31, P33, P34
		Efficiency and Time Saving	P2, P10, P12, P15, P21, P23, P25, P29, P30, P31
		Perceived Playfulness	P6, P15, P28

Themes	Sub-themes	Codes	Participants
		Emotional Engagement	P15, P21
		Attitude toward Using Generative AI	P1, P4, P10, P25, P27, P29, P31, P33
Concerns of Generative AI	Issues Specific to Technology Itself	Misinterpretations and Errors	P1, P4, P5, P6, P8, P10, P11, P12, P15, P16, P18, P24, P25, P27, P29, P33
		Challenges in Command Precision	P1, P6, P9, P10, P11, P12, P15, P17, P20, P24, P25, P26, P29, P33, P34
		Dependency of the Internet	P31, P28
		Ethical considerations	P31
	Issues Related to Research Experience	Autonomy Erosion and Skill Limitation	P10, P31
		AI Potential vs. Human Essence	P20
TPACK Insights	Technology Knowledge	Generative AI Skill Enhancement	P2, P6, P7, P8, P9, P12, P25
	Content Knowledge	Content Knowledge Enhancement	P4
	Pedagogical Knowledge	Pedagogical Knowledge Enhancement	P4, P9, P26
	Techno-content Knowledge	Prompt Precision Advantage	P17, P19, P21, P24, P32, P33, P34, P35
		Need for Subject Mastery	P1, P4, P34
	Techno-pedagogical Knowledge	Technology- education Integration	P16, P7, P9, P14
Future Orientations		Behavioral Intention to Use	P2, P3, P4, P10, P12, P13, P14, P15, P16, P19, P21, P22, P25, P29, P31, P32, P33, P35

The thematic analysis of pre-service teachers' reflective thoughts on their experiences with generative AI for question creation has unveiled themes encapsulating Previous Experiences, Positive Aspects of Generative AI, Concerns of Generative AI, Technological Pedagogical and Content Knowledge (TPACK) Insights, and Future Orientations. This analysis elucidates the multifaceted nature of integrating generative AI into educational practices, highlighting the interplay between educators' prior experiences and their perceptions of AI's potential benefits and limitations. The reflections underscore the critical role of TPACK in effectively leveraging AI technologies for pedagogical purposes and shed light on the evolving perspectives of future educators toward the

adoption of generative AI tools in shaping innovative educational paradigms.

Previous experiences significantly influenced participants' engagement with generative AI, leading to a shift in their perception toward the technology. For instance, Participant 1 (P1) remarked, "Initially skeptical, I was astonished by the AI's ability to understand and execute complex question design commands," indicating a transformative experience from doubt to appreciation. Similarly, AI familiarity comfort was evident in P5's and P6's narratives, with P5 noting, "Having used AI tools before, I found the interface intuitive and the process less daunting." In contrast, AI unfamiliarity emerged as a barrier for P4 and P19, where P4 stated, "The unfamiliarity with AI was initially a hurdle, making the task seem more complicated than it was."

The analysis highlighted several positive aspects of generative AI, divided into the technology itself and the experience it facilitates. The perceived ease of use was frequently mentioned, with P6 and P33 praising AI's user-friendly interface. The dialogic interface was appreciated for enabling a conversational approach to question creation, as P22 described: "It felt like having a knowledgeable partner assisting me throughout the process." The diversity of options and language adaptability were also seen as significant advantages, enhancing the creativity and inclusivity of the question creation process (P1, P5, P12, P21).

The experiential benefits included perceived usefulness, with many participants (e.g., P2, P3, P20) acknowledging AI's role in streamlining educational content creation. P15 and P31 underscored efficiency and time saving, noting the reduction in time required to generate quality questions. Perceived playfulness and emotional engagement were highlighted by participants like P28 and P15, who found the interaction with AI enjoyable and engaging, potentially influencing their attitude toward using generative AI in educational contexts positively.

Despite the positive aspects, concerns were raised, particularly regarding misinterpretations and errors in AI-generated content. P1 and P24 experienced inconsistencies in the AI's understanding of their commands. For example, Participant 24 said, "I saw that the questions prepared with insufficient command were contradictory, because the artificial intelligence that receives the command can prepare as much as it understands. " The challenge in command precision was a recurring theme, with participants like P10 and P34 emphasizing the need for clear and precise instructions to obtain the desired outcomes. Participant 10 stated that at first, she was not satisfied with the content produced by Chat GPT, but then she elaborated her prompts and obtained more desired results. Concerns about dependency on internet access and ethical issues related to AI use in education were also mentioned by P28 and P31, highlighting the need for critical consideration of AI's role in educational settings. P31 claimed that "I have seen that the use of AI has negative as well as positive aspects and that it should be used in a controlled manner. Artificial intelligence affects people negatively in many aspects such as creativity, thinking differently, privacy, being an individual, thinking differently, addiction, and copying. " and with these statements, she stated that AI can cause problems in many aspects, especially ethics.

This study provided valuable insights into how engagement with generative AI could enhance TPACK. Participants reported improvements in technological knowledge, particularly in generative AI skill enhancement (P2, P6, P25), and content knowledge, with some noting how the experience enriched their understanding of the

subject matter. P4 and P9 mentioned pedagogical knowledge enhancement, recognizing the potential of AI in supporting diverse teaching strategies. For instance, P9 stated, "Through this process, I learned about artificial intelligence applications, I had a better understanding of Bloom's taxonomy, and I realized how difficult it is to prepare good and inclusive questions." The need for prompt precision and subject mastery was evident, as participants like P17 and P34 highlighted the importance of understanding both the content and technological tools to effectively integrate them into educational practices.

Participants' future orientations reflected a strong behavioral intention to use generative AI in their teaching practices, which was influenced by their experiences and perceived benefits. The readiness to embrace AI in education was marked by a recognition of its potential to transform educational content creation, enhance learning experiences, and support personalized learning paths. Upon examining the reflective narratives of the participants, it becomes evident that their intentions regarding the use of generative AI span across different stages of their educational and professional journey. A segment of the participants, notably Participant 21 and Participant 35, expressed a keen interest in employing generative AI tools during their tenure as students, underscoring the technology's potential to augment their learning and academic research capabilities. Conversely, another group, including P4, P14, P19, and P33, articulated a forward-looking perspective toward harnessing this technology in their forthcoming teaching careers, recognizing the significant benefits it could offer in crafting engaging and diverse educational content. For example, P33 said "It will be a tool that I will want to frequently use in my field, preschool. ..... It is a tool that children will have a lot of fun with, and they will definitely want to use it in the following years. Therefore, in the activities I will prepare, we can experience the use of technology with children." This situation demonstrates the pre-service teacher's willingness to use this new technology in the teaching profession. Furthermore, a few participants, such as P12 and P22, envisioned a dual application of generative AI, indicating their intention to leverage these advanced technological tools throughout their academic pursuits as students and subsequently as educators. This diverse range of intentions highlights the multifaceted appeal of generative AI across educational activities, from enhancing personal learning experiences to enriching the pedagogical process.

Discussion

Discussion on Pre and Post Training Evaluations

In the current study, the participants were tasked with evaluating the multiple-choice questions (MCQs) they generated both before and after undergoing training. The initial evaluation by pre-service teachers of the AI-generated MCQs concentrated on various factors, including the complexity of the items, relevance of the content, appropriateness for intended grade levels, and clarity of the items. Several participants perceived the questions to be overly simplistic, indicating a necessity for questions of medium or challenging difficulty. Research such as that by Han et al. (2023) and Agarwal et al. (2023) underscores the usefulness of AI in crafting educational content, albeit with a note of caution regarding possible inaccuracies. Conversely, studies by Kıyak et al. (2024) and Cheung et al. (2023) have showcased the capability of AI to produce MCQs of acceptable quality and efficiency. However, these studies provided highly detailed prompts for generating MCQs, including specific requirements regarding the type of item, difficulty level, number of options, and formulation of the item stem. This level of

detail in prompt-writing has led to differing outcomes. These observations emphasize the essential role of clear and precise prompt instructions in generating effective AI-produced content. They suggest that AI holds significant potential to revolutionize educational assessment by efficiently producing high-quality multiple-choice questions (MCQs). However, this potential is dependent on human oversight to guarantee both accuracy and contextual relevance.

The findings reveal a notable shift in the pre-service teachers' evaluation practices after training. Post-training evaluations of AI-generated multiple-choice questions were more detailed and focused, incorporating advanced criteria such as item type, option suitability, and learning outcome alignment. This reflects Indra et al.'s (2023) assertion on the necessity of human expertise in refining AI outputs for educational assessments. Stojanov (2023) and Su and Yang (2023) further emphasize the need for critical engagement and ethical considerations in the educational use of AI's. The evolution from superficial to nuanced evaluations underscores the essential role of domain expertise in effectively leveraging AI for educational content development, aligning with the findings of Zuckerman et al. (2023)'s on the efficiency of AI-assisted item creation. This progression suggests a promising integration pathway for AI in education, contingent on informed and deliberate application by educators.

The final evaluation of AI-generated multiple-choice questions revealed divergent viewpoints among the participants regarding the appropriateness and adequacy of options and distractors. Some participants appreciated the relevance and alignment of options generated by generative AI, while others criticized them for not adhering sufficiently to assessment principles. This feedback underscores the importance of crafting concise options, unambiguous, and directly related to the subject matter, reflecting participants' application of pedagogical insights. This dichotomy suggests a need for further refinement in AI's question generation process to better meet educational standards.

In the final evaluation phase, participants concentrated on assessing the validity, reliability, and utility of AI-generated questions. Seven participants specifically affirmed the content validity of these questions. This finding aligns with prior research highlighting AI's potential in grading and assessment tasks, as demonstrated by Kim et al. (2019) in their study on AI's alignment with human assessments, and by Terwiesch (2023), who emphasized ChatGPT's efficiency in educational contexts. However, it is crucial to acknowledge the limitations of this evaluation. The small sample size of participants providing feedback and the reliance on general statements may suggest a need for further investigation into how these participants understand and apply assessment principles in AI contexts. Future research should aim to delve deeper into this aspect to gain a more comprehensive understanding of the nuances involved in evaluating AI-generated educational content.

Reflecting on their experiences with generative AI in question generation, the participants' prior experiences significantly influenced their utilization of artificial intelligence (AI), with novices finding the process challenging and experienced users navigating it with ease. This variation underscores the impact of past experiences on technology adoption, emphasizing the necessity of customized support and training in educational settings to bridge the experience gap. This approach could cultivate a more inclusive environment where all individuals, regardless of their previous exposure to AI tools, can effectively engage with and benefit from such technologies

in educational contexts.

When the participants' views on the research process were analyzed, the positive and negative aspects of the participants' experiences were underlined. They appreciated the ease of use, conversational interface, speed in generating alternatives, and language flexibility. This result aligns with the broader educational discourse on technology acceptance (Davis, 1989). The perceived ease of use, conversational interface, and adaptability of language and options not only facilitate a more inclusive and creative educational content creation process but also foster a positive attitude toward the adoption of such technologies in pedagogical practices (Venkatesh & Davis, 2000). Studies such as Jishnu et al. (2023) and Abdullah and Zaid (2023) affirm the benefits of AI for academic purposes and emphasize the importance of training for its effective use. This underscores the dual nature of AI tools in education: while offering significant advantages in accessibility and efficiency, their optimal use requires user education to overcome challenges and leverage their full potential.

Reflecting on the integration of generative AI in educational contexts, participants highlighted its utility in facilitating various aspects of academic and personal life, emphasizing its role in saving time and effort. This perspective is supported by research indicating that AI tools, such as ChatGPT can significantly enhance the efficiency of educational processes, such as grading and content creation (Kim et al., 2019; Cheung et al., 2023). Moreover, studies have demonstrated the positive reception of AI among educators and students alike, attributing its acceptance to its perceived ease of use and usefulness (Zhang et al., 2023; Dwivedi et al., 2023). These findings collectively underscore the transformative potential of AI in education, suggesting its capacity to enhance learning outcomes and operational efficiency. However, the variation in perceptions of AI's utility highlights the importance of comprehensive training to optimize its benefits across educational settings.

The emotional engagement experienced by participants while using AI tools, marked by enjoyment and excitement, is echoed in studies by Jepkemoi et al. (2024), and Das and Madhusudan (2024), which highlight the educational and emotional benefits of ChatGPT. This engagement is crucial for adaptive learning experiences and motivation, suggesting that positive emotional responses to AI tools can significantly influence their continued use in educational settings. This finding underscores the importance of incorporating engaging and interactive AI technologies to enhance learning experiences, making them not only educational but also enjoyable.

Significant concerns related to misinterpretations, errors, dependency on internet access, and ethical considerations temper the enthusiasm for generative AI. These apprehensions underscore the critical need for a nuanced understanding and careful navigation of AI technologies. The emphasis on precise command execution, along with the risks of autonomy reduction and skill limitations, underscores the urgent necessity for educators to acquire a comprehensive understanding of AI tools. This aligns with research on AI "hallucination," where AI produces nonsensical or unfaithful content (Maynez et al., 2020; Heaven, 2020). Han et al. (2023) and Naher et al. (2023) emphasize the need for caution and expert oversight to mitigate these issues. Awareness of AI's limitations among educators is crucial as it influences the evaluation of AI-generated content and its integration into educational settings. This underscores the importance of critically engaging with AI outputs.

The challenges participants faced in writing commands for AI and the need for additional prompts to achieve desired outcomes underscore the importance of prompt writing in educational AI tools. As AI becomes integral to content creation and management, educators must develop skills in crafting effective prompts to ensure that AI-generated content meets pedagogical standards. This shift necessitates continuous learning and adaptation by educators to harness AI's potential for creating diverse and engaging learning experiences, highlighting the evolving nature of teaching roles in the integration of advanced technologies.

Reflecting on the dual nature of AI in education, concerns about its potential to supplant human roles and diminish creativity are prominent among some educators. This discourse, highlighted by García-Peñalvo (2023) and further debated by García-Peñalvo et al. (2021) and Naher et al. (2023), underscores the critical balance needed to integrate AI. The key lies not in eschewing AI but in harnessing its capabilities to bolster critical thinking and creativity, ensuring AI's role is complementary, enhancing the teaching and learning ecosystem without diminishing the human essence of education.

Enhancement of TPACK through engagement with generative AI offers a promising avenue for educators' professional development. The study's findings indicate that interaction with AI can enhance technological, content, and pedagogical knowledge, facilitating a more profound integration of technology in education. This enhancement is particularly evident in the reported improvements in AI skills, content knowledge, and pedagogical strategies. This enhancement aligns with Jang and Lee's research (2023) which emphasize the role of AI in teacher trainning and improving educational outcomes. It highlights the necessity for educators to adeptly navigate AI tools, enriching learning while addressing ethical concerns and pedagogical integrity, as supported by insights from Jaiswal and Arun (2021), Aldeman et al. (2021), and research emphasizing AI's potential in education (Coniam, 2008; Jeon & Lee, 2023; Kohnke et al., 2023).

The participants' inclination towards integrating generative AI into their future academic and teaching careers highlights a transformative perspective on educational technology. This reflects broader research trends, as highlighted by research articles (Chan & Hu, 2023; Das & J.V., 2024; Nyaaba and Zhai, 2024; Wang et al., 2023). These studies collectively demonstrate a growing eagerness among educators and students to embrace AI to enhance educational effectiveness and personalize learning experiences. These studies illustrate a consensus on the potential of AI to enhance teaching methodologies, academic research, and classroom practices. The strong behavioral intention to use generative AI, influenced by its perceived benefits, signals a shift toward more innovative, technology-infused educational environments. This echoes the call by Ertmer and Ottenbreit-Leftwich (2010) for educators to be not only consumers but also innovators and critical users of technology.

Implications

This study carries important implications for educational policy, teacher education, and curriculum development. Firstly, they highlight the need for policymakers to prioritize the integration of AI technologies within educational curricula. This requires equipping future teachers with the necessary skills to effectively utilize these tools to enhance student learning outcomes. Achieving this goal involves embedding AI-related content into teacher

education programs and offering continuous professional development opportunities for in-service teachers, enabling them to remain current with technological advancements.

Second, the study highlights that the outputs generated by generative AI may not always be accurate or effective. To enhance pre-service teachers' proficiency in using these tools more effectively, it is imperative to offer specialized training that combines prompt engineering with expertise in assessment and evaluation. This multidisciplinary approach can empower future educators to customize AI-generated content to meet specific educational objectives, ensuring that the integration of AI technologies into teaching and learning processes is both effective and pedagogically sound. Finally, the study suggests that curriculum developers and instructors should design learning experiences that enable pre-service teachers to engage hands-on with AI technologies. By integrating project-based learning assignments that require the application of AI tools in instructional design and assessment, teacher education programs can prepare future educators to effectively incorporate AI into their teaching practices. This integration enhances the learning experience for students in a rapidly evolving digital age.

Limitations and Future Studies

The limitations of this study include its focus on a specific group of pre-service teachers, which may restrict the generalizability of the findings to broader educational contexts. Additionally, the evolving nature of AI technology means that participants' experiences and perceptions could rapidly change as new developments emerge. Future research could expand by exploring diverse educational settings, incorporating longitudinal studies to track how AI integration in education evolves, and examining the impact of comprehensive AI training programs on teaching and learning outcomes.

Another limitation is related to the item type. The study's focus was limited to one type of question, excluding a comparative analysis across different assessment methods such as performance assignments, peer assessment, or self-assessment. Investigating educators' experiences with generative AI in various assessment practices could offer valuable insights into measuring 21st-century skills. While not specified in the question generation phase, participants exclusively chose ChatGPT and Google Bard, although this was not specified in the study's protocol. Consequently, the results obtained from this research can only be evaluated within the context of these specific technologies. This limitation underscores the potential variability in outcomes that may arise when using different generative AI platforms. To expand the scope of future research, a broader array of AI tools could be included. This approach would offer a deeper insight into the impact of various AI technologies on the creation of educational content and the implementation of assessment practices.

References

Abdullah, Z., & Zaid, N. M. (2023). Perception of Generative Artificial Intelligence in higher education research. *Innovative Teaching and Learning Journal*, *7*(2), 84-95.

Agyei, D. D., & Voogt, J. M. (2011). Exploring the potential of the will, skill, tool model in Ghana: Predicting prospective and practicing teachers' use of technology. *Computers & Education*, *56*(1), 91-100.

Aldeman, N. L. S., Aita, K., Machado, V. P., da Mata Sousa, L. C. D., Coelho, A. G. B., da Silva, A. S., Mendes, A. P. D., Neres, F. J. D., & do Monte, S. J. H. (2021). Smartpath (k): A platform for teaching glomerulopathies using machine learning. *BMC Medical Education*, *21*(1), 248-267. https://doi.org/10.1186/s12909-021-02680-1

Altun, H., & Usta, E. (2019). The effects of programming education planned with TPACK framework on learning outcomes. *Participatory Educational Research, 6*(2), 26-36.

Arkhipova, M. V., Belova, E. E., Gavrikova, Y. A., Pleskanyuk, T. N., & Arkhipov, A. N. (2019). *Reaching Generation Z. Attitude Toward Technology Among the Newest Generation of School Students*. In: Popkova, E., Ostrovskaya, V. (eds) Perspectives on the Use of New Information and Communication Technology (ICT) in the Modern Economy. Springer, Cham. https://doi.org/10.1007/978-3-319-90835-9_114

Aydin, Ö., & Karaarslan, E. (2023). Is chatgpt leading generative ai? What is beyond expectations?. *Academic Platform Journal of Engineering and Smart Systems*, *11*(3), 118-134. https://doi.org/10.21541/apjess.1293702

Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, *7*(1), 52-62.

Baker, T., & Smith, L. (2019). Educ-AI-tion rebooted? Exploring the future of artificial intelligence in schools and colleges. https://media.nesta.org.uk/documents/Future_of_AI_and_education_v5_WEB.pdf

Berg, C. (2023). The case for generative AI in scholarly practice. https://papers.ssrn.com/sol3/papers.cfm?abstract_id= 4407587

Bhowmick, A. K., Jagmohan, A., Vempaty, A., Dey, P., Hall, L., Hartman, J., ... & Maheshwari, H. (2023, November). Automating question generation from educational text. In International Conference on Innovative Techniques and Applications of Artificial Intelligence (pp. 437-450). Cham: Springer Nature Switzerland.

Bolukbasi, T., Chang, K. W., Zou, J. Y., Saligrama, V., & Kalai, A. T. (2016). Man is to computer programmer as woman is to homemaker? debiasing word embeddings. *Advances in Neural İnformation Processing Systems*, *29*, 1-9.

Bozkurt, A., Xiao, J., Lambert, S., Pazurek, A., Crompton, H., Koseoglu, S., Farrow, R., Bond, M., Nerantzi, C., Honeychurch, S., Bali, M., Dron, J., Mir, K., Stewart, B., Costello, E., Mason, J., Stracke, C. M., Romero-Hall, E., Koutropoulos, A., Toquero, C. M., Singh, L., Tlili, A., Lee, K., Nichols, M., Ossiannilsson, E., Brown, M., Irvine, V., Raffaghelli, J. E., Santos-Hermosa, G., Farrell, O., Adam, T., Thong, Y. L., Sani-Bozkurt, S., Sharma, R. C., Hrastinski, S., & Jandrić, P. (2023). Speculative futures on ChatGPT and generative Artificial Intelligence (AI): A collective reflection from the educational landscape. *Asian Journal of Distance Education*, *18*(1), 1-10. https://www.asianjde.com/ojs/index.php/AsianJDE/article/view/709

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77-101.

Cao, L. (2022). Ai in finance: Challenges, techniques, and opportunities. *ACM Computing Surveys (CSUR), 55*(3), 1-38.

Chan, C. K. Y., & Hu, W. (2023). *Students' voices on Generative AI: Perceptions, benefits, and challenges in higher education*. arXiv. https://arxiv.org/abs/2305.00290

Chan, C. K. Y., & Zhou, W. (2023). *Deconstructing student perceptions of Generative AI (GenAI) through an Expectancy Value Theory (EVT)-based instrument* [Preprint]. arXiv. https://arxiv.org/abs/2305.01186

Chan, K. S., & Zary, N. (2019). Applications and challenges of implementing artificial intelligence in medical education: Integrative review. *JMIR Medical Education*, *5*(1), 1-13. http://doi.org/10.2196/13930

Cheung, B. H. H., Lau, G. K. K., Wong, G. T. C., Lee, E. Y. P., Kulkarni, D., Seow, C. S., ... & Co, M. T. H. (2023). *ChatGPT versus human in generating medical graduate exam questions–An international prospective study*. MedRxiv. https://doi.org/10.1101/2023.05.13.23289943

Choi, S., Jang, Y., & Kim, H. (2023). Influence of pedagogical beliefs and perceived trust on teachers' acceptance of educational artificial intelligence tools. *International Journal of Human–Computer Interaction*, *39*(4), 910-922.

Chounta, I. A., Bardone, E., Raudsep, A., & Pedaste, M. (2022). Exploring teachers' perceptions of Artificial Intelligence as a tool to support their practice in Estonian K-12 education. *International Journal of Artificial Intelligence in Education*, *32*(3), 725-755.

Christensen, R. (2002). Effects of technology integration education on the attitudes of teachers and students. *Journal of Research on Technology in Education*, *34*(4), 411-433. http://doi.org/10.1080/15391523.2002.10782359

Coniam, D. (2008). Evaluating the language resources of chatbots for their potential in English as a second language. *ReCALL*, *20*(1), 98-116.

Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 1-12.

Creswell, J. W. (2012). *Educational research*. Pearson.

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

Cunningham-Nelson, S., Boles, W., Trouton, L., & Margerison, E. (2019). A review of chatbots in education: Practical steps forward. In *Educators becoming agents of change: Innovate, integrate, motivate—Proceedings of the 30th Annual Conference for the Australasian Association for Engineering Education* (pp. 299-306). Engineers Australia.

Dans, E. (2023). ChatGPT… y el miedo a la innovación. Enrique Dans. http://bit.ly/3ZGGMTb

Das, S. R., & Madhusudan, J. V. (2024). Perceptions of higher education students towards ChatGPT usage. *International Journal of Technology in Education (IJTE)*, *7*(1), 86-106. https://doi.org/10.46328/ijte.583

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of ınformation technology. *MIS Quarterly*, *13*(3), 319-340.

Ding, L., Li, T., Jiang, S., & Gapud, A. (2023). Students' perceptions of using ChatGPT in a physics class as a virtual tutor. *International Journal of Educational Technology in Higher Education*, *20*(1), 1-18. https://doi.org/10.1186/s41239-023-00434-1

Durall, E., & Kapros, E. (2020, July 19–24). *Co-design for a competency self-assessment chatbot and survey in science education* [Conference paper]. Learning and Collaboration Technologies. Human and

Technology Ecosystems: 7th International Conference, Copenhagen, Denmark (pp. 13-24).

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., ... & 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*, *71*, 1-63. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Ertmer, P. A., & Ottenbreit-Leftwich, A. T. (2010). Teacher technology change: How knowledge, confidence, beliefs, and culture intersect. *Journal of Research on Technology in Education*, *42*(3), 255-284.

Ertmer, P. A., Ottenbreit-Leftwich, A. T., Sadik, O., Sendurur, E., & Sendurur, P. (2012). Teacher beliefs and technology integration practices: A critical relationship. *Computers & Education*, *59*(2), 423-435. https://doi.org/10.1016/j.compedu.2012.02.001

García-Peñalvo, F. J., Corell, A., Abella-García, V., & Grande-de-Prado, M. (2021). Recommendations for mandatory online assessment in higher education during the COVID-19 pandemic. In D. Burgos, A. Tlili, & A. Tabacco (Eds.), *radical solutions for education in a crisis context. COVID-19 as an opportunity for global learning* (pp. 85-98). Springer Nature. https://doi.org/10.1007/978-981-15-7869-4_6

Gizchina.com. (2024, January 19). *OpenAI cooperates with Arizona State University to use GPT-4*. https://www.gizchina.com/2024/01/19/openai-cooperates-with-arizona-state-university-to-use-gpt-4/

Han, Z., Battaglia, F., Udaiyar, A., Fooks, A., & Terlecky, S. R. (2023). An explorative assessment of ChatGPT as an aid in medical education: Use it with caution. *Medical Teacher*, 1-17. htttp://doi.org/10.1080/0142159X.2023.2271159

Heaven, W. D. (2020, July 20). OpenAI's new language generator GPT-3 is shockingly good—and completely mindless. *MIT Technology Review*, https://www.technologyreview.com/2020/07/20/1005454/openaimachine-learning-language-generator-gpt-3-nlp/

Indran, I. R., Paramanathan, P., Gupta, N., & Mustafa, N. (2023). Twelve tips to leverage AI for efficient and effective medical question generation: A guide for educators using Chat GPT. *Medical Teacher*, 1-6.

Islam, I., & Islam, M. N. (2023). Opportunities and challenges of chatgpt in academia: A conceptual analysis. *Authorea Preprints*. https://doi.org/10.22541/ au.167712329.97543109/v1

Jaiswal, A., & Arun, C. J. (2021). Potential of artificial ıntelligence for transformation of the education system in India. *International Journal of Education and Development Using Information and Communication Technology*, *17*(1), 142-158.

Jang, M., & Lee, H. W. (2023). Pre-service teachers' education needs for aı-based education competency. *Educational Technology International*, *24*(2), 143-168.

Jeon, J., & Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies*, 1-20.

Jepkemoi, B., Mulwa, P. K., & Mwanda, S. O. (2024). Influence of ChatGPT affordances on adaptive learning experiences among undergraduate religious education teacher trainees at the University of Nairobi, Kenya. *Canadian Journal of Educational and Social Studies*, *4*(1), 25-35. https://doi.org/10.53103/cjess.v4i1.206

Jishnu, D., Srinivasan, M., Dhanunjay, G. S., & Shamala, R. (2023). Unveiling student motivations: A study of ChatGPT usage in education. *ShodhKosh: Journal of Visual and Performing Arts*, *4*(2), 65–73. https://doi.org/10.29121/shodhkosh.v4.i2.2023.503

John, P. (2005). The sacred and the profane: Subject sub-culture, pedagogical practice and teachers' perceptions of the classroom uses of ICT. *Educational Review*, *57*(4), 471–490. https://doi.org/10.1080/00131910500279577

Joo, Y. J., Park, S., & Lim, E. (2018). Factors influencing preservice teachers' intention to use technology: TPACK, teacher self-efficacy, and technology acceptance model. *Educational Technology & Society*, *21*(3), 48–59. https://www.jstor.org/stable/26458506

Jovanović, M. (2023). *Generative Artificial Intelligence: Trends and prospects*. https://www.computer.org/csdl/magazine/co/2022/10/09903869/1H0G6xvtREk. 0.1109/MC.2022.3192720.

Kadaruddin, K. (2023). Empowering education through generative AI: Innovative instructional strategies for tomorrow's learners. *International Journal of Business, Law, and Education*, *4*(2), 618-625.

Kaplan-Rakowski, R., Grotewold, K., Hartwick, P., & Papin, K. (2023). Generative AI and teachers' perspectives on its implementation in education. *Journal of Interactive Learning Research, 34*(2), 313-338.

Karsenti, T. (2019). Artificial intelligence in education: The urgent need to prepare teachers for tomorrow's schools. *Formation et Profession*, *27*(1), 105-111. http://dx.doi.org/10.18162/fp.2018.a166

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*, 102-274. https://doi.org/10.1016/j.lindif.2023.102274

Keleş, H. (2022). Artificial intelligence applications in medicine. *Journal of Kırıkkale University Faculty of Medicine*, *24*(3), 604-613. https://doi.org/10.24938/kutfd.1214512

Khan, I., Ahmad, A. R., Jabeur, N., & Mahdi, M. N. (2021). An artificial intelligence approach to monitor student performance and devise preventive measures. *Smart Learning Environments*, *8*(1), 1-18. https://doi.org/10.1186/s40561-021-00161-y

Kıyak, Y. S., Coşkun, Ö., Budakoğlu, I. İ., & Uluoğlu, C. (2024). ChatGPT for generating multiple-choice questions: Evidence on the use of artificial intelligence in automatic item generation for a rational pharmacotherapy exam. *European Journal of Clinical Pharmacology*, 1-7. https://doi.org/10.1007/s00228-024-03649-x

Kim, H. S., Kim, N. Y., & Cha, Y. (2021). Is it beneficial to use AI chatbots to improve learners' speaking performance? *Journal of ASIA TEFL*, *18*(1), 161-178. https://doi.org/10.18823/asiatefl.2021.18.1.10.161

Kim, M., Yi, S., & Lee, D. (2019). Between living and nonliving: Young children's animacy judgments and reasoning about humanoid robots. *PLOS One*, *14*(6), e0216869. https://doi.org/10.1371/journal. pone.0216869

Koehler, M., & Mishra, P. (2009). What is technological pedagogical content knowledge (TPACK)?. *Contemporary Issues in Technology and Teacher Education, 9*(1), 60-70.

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). Exploring generative artificial intelligence preparedness among university language instructors: A case study. *Computers and Education: Artificial Intelligence*, *5*, 100-156.

Korn, J., & Kelly, S. (2023). *New York City public schools ban access to AI tool that could help students cheat*. CNN Business.

Korngiebel, D. M., & Mooney, S. D. (2021). Considering the possibilities and pitfalls of Generative Pre-trained

Transformer 3 (GPT-3) in healthcare delivery. *NPJ Digital Medicine*, *4*(1), 1-3. https://doi.org/10.1038/s41746-021-00464-x

Kuhail, M. A., Alturki, N., Alramlawi, S., & Alhejori, K. (2023). Interacting with educational chatbots: A systematic review. *Education and Information Technologies*, *28*(1), 973-1018.

Kuleto, V., Ilić, M. P., Bucea-Manea-Ţoniş, R., Ciocodeică, D. F., Mihălcescu, H., & Mindrescu, V. (2022). The attitudes of K–12 schools' teachers in Serbia towards the potential of artificial intelligence. *Sustainability*, *14*(14), 1-15. https://doi.org/10.3390/su14148636

Lambert, J. L., & Lane, S. C. (2004). Technology integration expertise among middle school teachers. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *48*(3), 463-467. https://doi.org/10.1177/154193120404800340

Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education*, *21*(2), 1-13. https://doi.org/10.1016/j.ijme.2023.100790

Lin, H. (2022). Influences of Artificial Intelligence in education on teaching effectiveness: The mediating effect of teachers' perceptions of educational technology. *International Journal of Emerging Technologies in Learning (Online)*, *17*(24), 144-156. https://doi.org/10.3991/ijet.v17i24.36037

Liu, S. H. (2011). Factors related to pedagogical beliefs of teachers and technology integration. *Computers & Education*, *56*(4), 1012-1022. https://doi.org/10.1016/j.compedu.2010.12.001

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, *13*(4), 410-420. https://doi.org/10.3390/educsci13040410

Łodzikowski, K., Foltz, P. W., & Behrens, J. T. (2023). Generative AI and its educational implications. https://arxiv.org/ftp/arxiv/papers/2401/2401.08659.pdf

Maynez, J., Narayan, S., Bohnet, B., & McDonald, R. (2020). On faithfulness and factuality in abstractive summarization. *arXiv.* https://arxiv.org/pdf/2005.00661.pdf

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (1955). A proposal for the Dartmouth summer research project on artificial intelligence. http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html

Mello, R. F., Freitas, E., Pereira, F. D., Cabral, L., Tedesco, P., & Ramalho, G. (2023). Education in the age of Generative AI: Context and recent developments. https://arxiv.org/pdf/2309.15004.pdf

Mizumoto, A., & Eguchi, M. (2023). Exploring the potential of using an AI language model for automated essay scoring. *Research Methods in Applied Linguistics*, *2*(2), 1-13. https://doi.org/10.1016/j.rmal.2023.100050

Naher, J., Maruf, M. H., Bakht, S. T., & Sadaf, S. S. (2023). A survey to understand the experience of chatgpt usage among engineering university students in Bangladesh. *Technium*, *15*, 1-8.

Nair, I., & Das, V. M. (2012). Using Technology Acceptance Model to assess teachers' attitude towards use of technology as teaching tool: A SEM Approach. *International Journal of Computer Applications*, *42*(2), 1-6.

Nyaaba, M., & Zhai, X. (2024). Generative AI professional development needs for teacher educators. *Journal of AI*, *8*(1), 1-13. https://doi.org/10.61969/jai.1385915

Okonkwo, C. W., & Ade-Ibijola, A. (2020). Python-Bot: A chatbot for teaching python programming. *Engineering Letters*, *29*(1) 25-34. https://eds.p.ebscohost.com/eds/pdfviewer/pdfviewer?vid=0&sid=00b299bf-2e83-4e9d-baf4-2cf9227b5c5a%40redis

Pavlik, J. V. (2023). Collaborating with ChatGPT: Considering the implications of generative artificial intelligence for journalism and media education. *Journalism & Mass Communication Educator*, *78*(1), 84-93. https://doi.org/10.1177/10776958221149577

Qadir, J. (2023, May). Engineering education in the era of ChatGPT: Promise and pitfalls of generative AI for education. In *2023 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1-9). IEEE.

Ribeiro, J., Lima, R., Eckhardt, T., & Paiva, S. (2021). Robotic process automation and artificial intelligence in industry 4.0–a literature review. *Procedia Computer Science, 181*, 51-58.

Sengupta, S., & Chakraborty, T. (2020). Use of chatbots in higher education: A study of student engagement and satisfaction. *Education and Information Technologies*, *25*(6), 5147-5165. https://doi.org/10.1007/s10639-020-10171-x

Shaheen, M. Y. (2021). Applications of Artificial Intelligence (AI) in healthcare: A review. *Science Open Preprints.* https://doi.org/10.14293/S2199-1006.1.SOR-.PPVRY8K.v1

Shaikh, F., Dehmeshki, J., Bisdas, S., Roettger-Dupont, D., Kubassova, O., Aziz, M., & Awan, O. (2021). Artificial intelligence-based clinical decision support systems using advanced medical imaging and radiomics. *Current Problems in Diagnostic Radiology*, *50*(2), 262-267. https://doi.org/10.1067/j.cpradiol.2020.05.006

Shen-Berro, J. (2023). *New York City Schools blocked ChatGPT. Here's what other large districts are doing.* Chalkbeat.

Stojanov, A. (2023). Learning with ChatGPT 3.5 as a more knowledgeable other: An autoethnographic study. *International Journal of Educational Technology in Higher Education*, *20*(1), 1-17.

Su, J., & Yang, W. (2023). Unlocking the power of ChatGPT: A framework for applying generative AI in education. *ECNU Review of Education, 6*(3), 355-366. https://doi.org/10.1177/20965311231168423.

Tate, T., Doroudi, S., Ritchie, D., & Xu, Y. (2023). Educational research and AI-generated writing: Confronting the coming tsunami. https://osf.io/preprints/edarxiv/4mec3

Teo, T. (2011). Factors influencing teachers' intention to use technology: Model development and test. *Computers & Education*, *57*(4), 2432–2440

Teo, T. (2012). Examining the intention to use technology among pre-service teachers: An integration of the Technology Acceptance Model and Theory of Planned Behavior. *Interactive Learning Environments*, *20*(1), 3-18. https://doi.org/10.1080/10494821003714632

Teo, T. & Zhou, M. (2017). The influence of teachers' conceptions of teaching and learning on their technology acceptance. *Interactive Learning Environments*, *25*(4), 513-527. https://doi.org/10.1080/10494820.2016.1143844

Terwiesch, C. (2023). Would Chat GPT3 get a Wharton MBA? A prediction based on its performance in the operations management course. *Mack Institute for Innovation Management at the Wharton School, University of Pennsylvania*. https://mackinstitute.wharton.upenn.edu/wp-content/uploads/2023/01/Would-ChatGPT-get-a-Wharton-MBA.pdf

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), 1-23. https://doi.org/10.1186/s40561-023-00237-x

Turing, A. (1950). Machines informatiques et intelligence. *Mind*, *49*, 433-460.

Tzirides, A. O., Zapata, G., Searsmith, D., Saini, A., Cope, B., Kalantzis, M., ... &Kastania, P. (2023). Generative AI: Implications and applications for education. *Article In Review*.

https://doi.org/10.48550/arXiv.2305.07605

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the Technology Acceptance Model: Four longitudinal field studies. *Management Science*, *46*(2), 186-204.

Villegas-Ch, W., Sánchez-Viteri, S., & Roman-Cañizares, M. (2021). Academic activities recommendation system for sustainable education in the age of COVID-19. *Informatics*, *8*(2), 29-39. http://dx.doi.org/10.3390/informatics8020029

Wang, S., Sun, Z., & Chen, Y. (2023). Effects of higher education institutes' artificial intelligence capability on students' self-efficacy, creativity and learning performance. *Education and Information Technologies*, *28*(5), 4919-4939.

Wetzel, K., & Marshall, S. (2011). TPACK goes to sixth grade: Lessons from a middle school teacher in a high-technology-access classroom. *Journal of Digital Learning in Teacher Education, 28*(2), 73-81.

Xia, Q., Chiu, T. K. F, Lee, M., Temitayo I., Dai, Y., & Chai, C. S. (2022). A self-determination theory design approach for inclusive and diverse artificial intelligence (AI) K-12 education. *Computers & Education*, *189*, 1-15. https://doi.org/10.1016/j.compedu.2022.104582

Yang, T. C., & Chen, J. H. (2023). Pre-service teachers' perceptions and intentions regarding the use of chatbots through statistical and lag sequential analysis. *Computers and Education: Artificial Intelligence*, *4*, 100-119. https://doi.org/10.1016/j.caeai.2022.100119

Zastudil, C., Rogalska, M., Kapp, C., Vaughn, J., & MacNeil, S. (2023). Generative AI in computing education: Perspectives of students and instructors. https://arxiv.org/pdf/2308.04309.pdf

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

Zhang, C., Schießl, J., Plößl, L., Hofmann, F., & Gläser-Zikuda, M. (2023). Acceptance of artificial intelligence among pre-service teachers: A multigroup analysis. *International Journal of Educational Technology in Higher Education, 20*(1), 49-65.

Zhang, K., & Aslan, A. B. (2021). AI technologies for education: Recent research & future directions. *Computers and Education: Artificial Intelligence*, *2*, 1-11. https://doi.org/10.1016/j.caeai.2021.100025

Zheng, L., Niu, J., Zhong, L., & Gyasi, J. F. (2023). The effectiveness of artificial intelligence on learning achievement and learning perception: A meta-analysis. *Interactive Learning Environments*, *31*(9), 5650-5664. https://doi.org/10.1080/10494820.2021.2015693

Zuckerman, M., Flood, R., Tan, R. J., Kelp, N., Ecker, D. J., Menke, J., & Lockspeiser, T. (2023). ChatGPT for assessment writing. *Medical Teacher*, *45*(11), 1224-1227. https://doi.org/10.1080/0142159X.2023.2249239

Author Information

Kübra Karakaya Özer

http://orcid.org/0000-0002-0208-7870

Eskişehir Osmangazi University

Turkiye

Contact e-mail: *karakayakubra@gmail.com*