IJEMST International Journal of Education in Mathematics, Science and Technology

www.ijemst.net

Generative Artificial Intelligence (GenAI) Meets Assessment: Experimental Insights into Teacher Candidates' Attitudes and Acceptance

Kübra Karakaya Özyer [1] *, Betül Aydın [2]

[1] Eskişehir Osmangazi University, Odunpazarı/Eskişehir, Turkiye, 0000-0002-0208-7870
[2] Eskişehir Osmangazi University, Odunpazarı/Eskişehir, Turkiye, 0000-0002-4739-6503
* Corresponding author: Kübra Karakaya Özyer (kozyer@ogu.edu.tr)

Article Info

Article History

Received:
28 June 2025

Revised:
7 October 2025

Accepted:
21 November 2025

Published:
1 January 2026

Keywords

Generative Artificial Intelligence
Assessment
Pre-service teachers

Abstract

This experimental study examined the impact of pre-service teachers' development of performance tasks utilizing differentiated generative artificial intelligence tools—textual, visual, and video-based—on their attitudes towards generative artificial intelligence (AI) and their acceptance levels of AI. A quasi-experimental approach was utilized to address the research issues, incorporating one control group and two randomly allocated experimental groups. Ninety-four pre-service teachers voluntarily participated in the study. The groups were categorized based on three distinct generative artificial intelligence tools utilized by the students in the development of a performance task. As a results of analyses, a significant increase in both positive attitude and acceptance level was reported in the Control group using text-based GenAI; an increase only in positive attitude in the Experiment 1 group; and a decrease in negative attitude in the Experiment 2 group. Engagement with differentiated generative AI tools resulted in notable alterations in in-group attitudes and levels of acceptability. This study underscores the essential requirement for pre-service teachers to implement a nuanced and diversified strategy for AI integration, acknowledging the advantages and limitations of different AI tools and their capacity to affect pre-service teachers' attitudes and acceptance in various manners.

Citation: Karakaya Özyer, K. & Aydın, B. (2026). Generative Artificial Intelligence (GenAI) meets assessment: Experimental insights into teacher candidates' attitudes and acceptance. *International Journal of Education in Mathematics, Science and Technology (IJEMST), 14*(1), 21-45. https://doi.org/10.46328/ijemst.5190

Introduction

The widespread adoption of GenAI tools has also resonated within the education community, rapidly becoming an integral part of educational processes. While universities strive to develop policies and guidelines concerning the use of these tools around the world (Crawford et al., 2023; Driessens & Pischetola, 2024; Moorehouse et al., 2023), a growing body of research focuses on guiding students toward the ethical and effective utilization of such technologies (Kadaruddin, 2023; Qadir, 2023). A crucial concern emerges regarding educators' capacity to effectively integrate these emerging technologies into their practices. Teachers' ability to synthesize technological, content, and pedagogical knowledge is paramount to their teaching effectiveness (Bugti et al., 2024). Therefore, enhancing pre-service teachers' competencies during teacher education programs is essential. Equipping them with the skills to leverage GenAI tools responsibly and effectively, while fostering awareness of ethical considerations, is crucial for preparing them to navigate the evolving landscape of education and effectively utilize these tools in their future teaching practices.

Recent studies emphasize the potential of GenAI tools to transform various stages of educational processes. These tools can rapidly analyze students' individual needs and generate tailored content, supporting personalized learning pathways (Owan et al., 2023; Su & Yang, 2023). GenAI can also serve as a digital assistant for teachers, efficiently addressing students' queries and fostering more interactive classroom environments (Aksu-Dünya & Yıldız-Durak, 2024). Furthermore, such tools can monitor and evaluate students' progress, providing educators with valuable insights for continuous improvement (Chiu et al., 2023; Westera et al., 2020).

The integration of GenAI in education not only aims to alleviate teachers' workload but also contributes to the creation of personalized and adaptive learning environments for students. By fostering more engaging and interactive educational settings, these tools encourage active student participation in their learning journey, enhancing their overall academic experiences (Baidoo-Anu & Ansah, 2023; Labadze et al., 2023; Santos, 2023). Moreover, research suggests that GenAI-supported applications can bolster students' problem-solving skills and nurture creative thinking processes, essential competencies in contemporary education (Luckin et al., 2016; Ou et al., 2024).

In addition to its benefits for students, the use of GenAI tools has been shown to positively impact teachers' job performance and pedagogical skills. Teachers can utilize these tools to critically evaluate their instructional practices and receive constructive feedback for improvement (Gunawan et al., 2021; Jaiswal & Arun, 2021; Owan et al., 2023). Notably, Lin (2022) demonstrated that AI-assisted teaching significantly enhances instructional effectiveness, underscoring the transformative potential of AI in education. Furthermore, studies on the integration of GenAI tools in the professional development of pre-service teachers suggest that these technologies can effectively support their growth and readiness for teaching (Lu et al., 2024). This evidence suggests that GenAI tools hold significant potential not only for transforming student learning experiences but also for empowering educators (or future educators) to refine their practices and maximize their impact within the educational landscape.

Exploring the Impact of Generative AI on Educational Assessment and Evaluation

The integration of Generative AI (GenAI) into educational assessment and evaluation has transformed key processes, including the development, administration, and evaluation of assessment tools. GenAI technologies, such as ChatGPT, have been shown to significantly reduce the time and effort required for tasks like question generation, enabling educators to focus on higher-order educational activities (Terwiesch, 2023). Moreover, in computer-adaptive testing (CAT), AI enhances item selection by dynamically tailoring questions to the examinee's ability level, utilizing advanced deep learning techniques (Mujtaba & Mahapatra, 2020; Liu et al., 2024). During evaluation, AI tools automate tasks such as scoring closed-ended questions, assessing open-ended responses, and generating comprehensive feedback, as seen in platforms like Gradescope (Latif & Zhai, 2024; Owan et al., 2023). AI-driven automated essay scoring systems ensure efficiency, scalability, and consistency in evaluating written work, while also supporting the development of rubrics to enhance fairness and accuracy in subjective assessments (Conrad & Rees, 2024; Yancey et al., 2023). These advancements underscore the transformative potential of GenAI in streamlining and enhancing educational assessment practices especially in the subjective assessment methods.

One prominent subjective assessment method for evaluating higher-order thinking skills is performance-based tasks. Performance tasks, which target students' cognitive skills like analysis, synthesis, and evaluation in Bloom's taxonomy, offer valuable opportunities to assess both the product and the process of student learning (Berk, 1986; Conrad & Rees, 2024; Palm, 2008). However, despite their potential for rich insights into students' learning, the development of performance tasks is often labor-intensive and demands considerable expertise, potentially discouraging teachers from utilizing them (Halagatti et al., 2023).Challenges also arise during the assessment stage, where teachers often encounter difficulties in objectively scoring performance tasks (Metin, 2013). Due to the nature of these tasks, assessment can rely heavily on subjective judgments, emphasizing the need for well-structured rubrics to maintain scoring consistency. Yet, creating an effective rubric and implementing it fairly for every student requires substantial effort and time (Metin, 2013). It is precisely in these areas of performance-based assessment that GenAI tools offer promising potential for support.

GenAI offers educators practical support in developing, organizing, and evaluating performance-based tasks, which are crucial for assessing higher-order thinking skills. GenAI tools, such as ChatGPT, Claude, and Gemini, empower educators to generate performance tasks aligned with specific learning objectives by leveraging well-structured prompts (Conrad & Rees, 2024). This streamlined development process helps to overcome the traditionally labor-intensive nature of creating performance tasks, a barrier that may have previously discouraged teachers from employing this valuable assessment method (Halagatti et al., 2023).Moreover, AI-driven platforms like Midjourney, DALL-E, Fliki.ai, and Flux.ai provide educators with the ability to create engaging visual materials to complement performance tasks, further enriching the assessment experience. In the evaluation phase, text-based GAI tools facilitate the development of analytic or holistic rubrics, ensuring alignment with predetermined assessment criteria (Conrad & Rees, 2024). Platforms such as Gradescope and ExamScore streamline the evaluation process by enabling automated assessment of student submissions based on teacher-designed rubrics. This automation significantly reduces educators' workloads while ensuring that assessments are personalized and reflective of each student's learning trajectory. The integration of GenAI in performance-based assessment not only addresses the challenges of rubric development and consistent scoring (Metin, 2013) but also supports educators in providing more effective and personalized feedback to students.

As discussed, GenAI tools offer significant benefits to teachers in the development and evaluation of performance-based tasks. These benefits, however, hinge on teachers' positive perceptions, attitudes, and behaviors toward these technologies. While the potential of AI in education is substantial, its successful integration is not without challenges. It requires careful consideration of several key factors, including teachers’ perceptions and attitudes toward these technologies, their willingness to embrace and adapt to AI-driven innovations, and their capacity to sustain creativity and pedagogical authenticity while leveraging AI in their instructional practices (Chung, 2024). Addressing these challenges is paramount for ensuring that AI tools are used effectively and ethically to enhance teaching and learning. The successful implementation of AI in education requires a holistic approach that considers not only the technological aspects but also the human factors that influence its adoption and impact.

Teachers’ Attitudes Towards Gen AI Technology in Learning Environments

Attitudes towards artificial intelligence (AI) are multifaceted, encompassing emotional, cognitive, and behavioral responses to these technologies. Ajzen and Fishbein's Theory of Planned Behavior (TPB) provides a valuable framework for understanding the development of these attitudes (Ajzen & Fishbein, 2000). The TPB suggests that an individual's intention to engage in a particular behavior, such as using AI tools, is shaped by three primary factors: attitude towards the behavior, subjective norms, and perceived behavioral control (Ajzen & Fishbein, 2000). The TPB highlights the interconnectedness of these factors in shaping behavioral intentions. For example, a teacher who holds a positive attitude toward AI, perceives supportive norms from their colleagues, and feels confident in their ability to use AI tools is more likely to integrate these technologies into their teaching practices. Therefore, understanding attitudes towards AI through a comprehensive theoretical lens like the TPB is crucial. These attitudes not only mirror opinions of the benefits and drawbacks of AI but also elucidate the elements that affect the acceptance and integration of AI in educational environments. Examining issues like educators' perceptions of AI's advantages and their perceived control over its deployment can facilitate the effective and ethical utilization of AI to improve teaching and learning.

Given the importance of teachers' attitudes in shaping their intention to use AI, it is crucial to examine current research on these attitudes. A review of recent studies reveals predominantly positive perceptions of GenAI tools among teachers and preservice teachers (Bezjak, 2024; Evmenova et al., 2024; Kim & Kim, 2024; Lee et al., 2024; Nyaaba et al., 2024; Pettersson et al., 2024). Many educators acknowledge the potential of these tools to enhance their professional work, specifically by reducing administrative workload and improving instructional quality through opportunities for self-reflection and gaining new perspectives (Chiu, 2023). This recognition aligns with the TPB, as teachers' positive attitudes toward GenAI stem from their perception of its benefits, such as improving learning outcomes. For example, ChatGPT can help educators brainstorm ideas, create rubrics, and develop learning questions, ultimately fostering student self-assessment and engagement (Chiu, 2023; Xia et al., 2024). Supporting this positive outlook, a study by Bower et al. (2024) involving 318 teachers across diverse disciplines found that most participants perceived GenAI as significantly impacting educational assessment. However, these perceptions varied depending on factors like years of teaching experience, gender, and the teachers' specific discipline (Bower et al., 2024).

Despite the predominantly positive perceptions of GenAI, teachers also voice significant concerns regarding its use in education (Chung, 2024; Kim & Kim, 2024; Lee et al., 2024). One prominent concern revolves around the potential negative impact of excessive AI reliance on students' critical thinking and active engagement in learning (Picton & Clark, 2024; Echave et al., 2024). For instance, over 50% of teachers in one study expressed fears that GenAI might hinder students' capacity for independent thought.

Further concerns center around data security issues related to student information entered into AI tools, the potential for AI to generate inaccurate content, and the risk of fostering overdependence on these technologies (Bezjak, 2024; Picton & Clark, 2024; Kim & Kim, 2024). Notably, a study by Kim and Kim (2024) involving primary and secondary school teachers emphasized concerns about misinformation generated by GenAI and advocated for the implementation of training programs to promote the effective and ethical use of AI tools. Such training could address teachers' concerns about AI, enhance their understanding of its capabilities and limitations, and empower them to leverage GenAI responsibly and effectively in their teaching practices. By acquiring the necessary knowledge and skills, teachers can develop positive attitudes towards AI and translate those attitudes into informed and beneficial behaviors, ultimately promoting successful AI integration in education.

Teachers' attitudes, whether positive or negative, are intrinsically linked to their acceptance of specific technologies. Research consistently demonstrates that as teachers' positive attitudes toward technology increase, so does their level of acceptance, ultimately leading to a greater integration of these technologies into their teaching practices (Hamzah et al., 2024). The Technology Acceptance Model (TAM), developed by Davis (1989), provides a valuable framework for understanding this relationship between attitudes and technology acceptance among educators (Nair & Das, 2012). At the core of TAM are three key constructs: perceived ease of use, perceived usefulness, and attitudes towards using technology (Davis, 1989).

Perceived ease of use centers on teachers' beliefs about the effortlessness of using technological tools. In contrast, perceived usefulness focuses on educators' expectations regarding how these technologies can enhance the effectiveness and quality of their teaching practices (Nair & Das, 2012). According to TAM, individuals' intentions to use technology are shaped by their perceptions of the technology and their attitudes towards using it. Similarly, Teo (2009) aimed to extend the original TAM by including additional variables such as computer self-efficacy, technological complexity and facilitating conditions to provide a more comprehensive model of pre-service teachers' acceptance of technology. In these comprehensive models, attitudes towards technology were demonstrated to be a direct and substantial predictor of behavioral intention. This finding emphasizes the importance of mindset in their willingness to adopt and incorporate technology into teaching practices.

Research specifically exploring the acceptance of GenAI in education has revealed a significant willingness among teachers to integrate these tools into their pedagogical practices (Arguson et al., 2023; Cabero-Almenara et al., 2024; Ofosu-Ampong, 2024). This positive trend aligns with the broader concept of technology acceptance, as highlighted by the Technology Acceptance Model (TAM), which posits that positive attitudes toward technology drive its adoption (Davis, 1989; Nair & Das, 2012). Studies consistently demonstrate that positive attitudes toward GenAI are a strong predictor of its adoption by educators. For instance, Ofosu-Ampong (2024) found that 84% of educators participating in their study expressed acceptance of GenAI technologies, with attitudes toward AI emerging as a key factor influencing this acceptance.

Similarly, Lu et al. (2024) identified a positive correlation between improved attitudes toward GenAI and increased intentions to utilize these technologies in education. The findings underscore that fostering favorable opinions of GenAI among educators can greatly enhance its effective incorporation into classrooms, as attitudes influence both the readiness to utilize and the actual usage behaviors (Hamzah et al., 2024). Consequently, the study highlights the crucial influence of teacher attitudes in facilitating the implementation of GenAI technologies. This underscores the necessity for focused techniques and professional development programs that assist educators in cultivating a constructive and informed comprehension of GenAI and its capacity to improve teaching and learning.

This study addresses critical gaps in understanding the impact of GenAI on pre-service teachers. While research explores educators' attitudes and acceptance of AI (e.g., Baidoo-Anu & Ansah, 2023; Bannister et al., 2023; Bozkurt, 2023), several limitations exist. First, existing studies often involve educators with limited AI experience (e.g., Ciftci et al., 2024; Lee et al., 2024; Pettersson et al., 2024), failing to capture the full potential of AI integration, particularly in cognitively demanding tasks like performance assessment development. Second, despite some evidence of positive effects of AI on attitudes and creativity (e.g., Cabero-Almenara et al., 2024; Lu et al., 2024; Ofosu-Ampong, 2024), inconsistent findings and a focus on higher education instructors rather than pre-service teachers necessitate further investigation. Third, there is a lack of experimental research specifically designed to *enhance* attitudes and acceptance, especially research that examines the interplay between these factors and creativity (e.g., Chandrasekera et al., 2024; Kim, 2023; Yin et al., 2024). This study directly addresses these gaps by experimentally investigating how using GenAI to develop performance tasks influences pre-service teachers' attitudes, acceptance, *and* creativity. By focusing on pre-service teachers, this research offers crucial insights for teacher preparation programs, contributing to more effective integration of AI in K-12 education, as highlighted by mapping reviews (e.g., Yusuf et al., 2024).

Research Aim

This study aims to investigate the impact of using GenAI tools in the development of performance tasks on pre-service teachers' attitudes towards AI and their acceptance of AI technology. In this context, answers were sought to the following research questions:

1. How does the use of GenAI tools for performance task development affect pre-service teachers' attitudes towards AI?
2. How does the use of GenAI tools for performance task development affect pre-service teachers' AI acceptance levels?

Methodology

Research Design

This study examined the impact of pre-service teachers' development of performance challenges utilizing GenAI tools on their views towards GenAI and their acceptance of AI. A quasi-experimental design was favored in this study. Quasi-experimental designs are studies that compare experimental and control groups to assess the impact of interventions, without employing random assignment (Creswell & Creswell, 2020). This study employed a quasi-experimental design, deemed suitable due to the selection of distinct courses as experimental and control groups in the Measurement and Evaluation course at the Faculty of Education.

Study Group

The study participants were third-year university students enrolled in the Faculty of Education. Three distinct groups were established for the research: one control group and two experimental groups. The groups were classified based on the various GenAI tools utilized by the students in the construction of the performance task (see Table 1).

Table 1. Participants, Groups, and Characteristics

Used GenAI	Gender	Control group	Experiment 1	Experiment 2
		ChatGPT	Canva (DALL-E)	Fliki.ai
Gender	Female	24	27	22
	Male	4	9	8

The pre-service teachers in the control group were instructed to use only ChatGPT, a text-based GenAI tool. In the first experimental group, participants were asked to use both ChatGPT and the visual GenI tool DALL-E. The pre-service teachers in the second experimental group were instructed to use Fliki.ai, an AI tool that generates videos in conjunction with ChatGPT when creating performance tasks. All students participating in the study were in Year 3, and these students were selected because they had taken the measurement and assessment course at this level. This ensured that the students had the necessary knowledge of measurement and assessment in the process of creating performance tasks.

Random assignments were used to determine the groups, but not to assign the participants to the groups. Within the framework of the research, three different sections of Year 3 were designated as A, B and C, and these classes were randomly selected to be assigned as control or experimental groups. At the beginning of the study there were 36 participants in the control group, 51 in the first experimental group and 52 in the second experimental group. However, in determining the final participants in the study, those who did not take the pre-test, did not complete the online training and did not take the post-test were excluded from the analysis. After this elimination process, the results of 28 pre-service teachers in the control group, 36 pre-service teachers in the first experimental group and 30 pre-service teachers in the second experimental group were analyzed.

Data Collection

Demographic Information Form

This form was used to collect basic demographic information such as the participants' gender, student number and the branch of the course they were taking.

General Attitude Scale towards Artificial Intelligence

It was developed by Schepman and Rodway (2023) in order to measure the participants' general attitudes towards AI. This scale was adapted into Turkish by Kaya et al. (2022). The scale consists of 20 items and two dimensions (positive attitude and negative attitude). Participants rate each item between 1 and 5. Total scores for positive and negative attitude dimensions were calculated separately. Cronbach's alpha internal consistency coefficients were calculated as .82 for the positive attitude dimension and .84 for the negative attitude dimension (Kaya et al., 2022).

Generative Artificial Intelligence Acceptance Scale

It was developed by Karacaoğlan-Yılmaz et al. (2023) to measure the level of acceptance of GenAI tools by individuals. The scale consists of 20 items and four sub-dimensions (performance expectation, effort expectation, facilitating conditions and social influence). Participants rated the scale items between 1 and 5. The highest score that can be obtained from the scale is 100 and the lowest score is 20. The evidence for the reliability of this scale was calculated as .97 with Cronbach's alpha and .95 with the test-retest method (Karacaoğlan-Yılmaz et al., 2023).

Procedure

Participants were randomly assigned to the control and experimental groups. Informed consent forms were distributed to the participants for their voluntary participation in the study and those who agreed to participate were included in the study and pre-tests were administered (see Figure 1). Then, a two-hour face-to-face training was given on creating a performance task and preparing a rubric. In this training, the definition, characteristics, usage areas and types of performance tasks were explained in detail and sample performance tasks were presented. In addition, the types, characteristics and usage examples of rubrics required for the evaluation of performance tasks were explained to the participants. Following this face-to-face training, the participants received an online and asynchronous training on GenAI. This training consisted of two parts. The first part aims to increase the general knowledge level of all participants about GenAI. The training covered the main purposes and benefits of GenAI, prompt engineering and how to create products with GenAI.

In the second part of the training, the trainings received by the groups were designed with different contents (see Table 2). The control group received only a training on the use of ChatGPT. Experiment 1 group received training on the use of DALL-E tool in addition to ChatGPT, and Experiment 2 group received training on Fliki.ai, a video creation tool, in addition to ChatGPT. These trainings were conducted online using the Edpuzzle platform. The Edpuzzle platform was preferred because it offers the opportunity to track how long the participants watch the given videos and to increase the attention and interaction of the viewers by adding questions on the videos.

Table 2. Details about Experiment Process

	Control Group	Experiment 1	Experiment 2
Pre-test	Demographic Information Form	Demographic Information Form	Demographic Information Form
	General Attitude towards AI	General Attitude towards AI	General Attitude towards AI
	GenAI Acceptance Scale	GenAI Acceptance Scale	GenAI Acceptance Scale
Process 1 (2 hours)	Performance Task and Rubric Training	Performance Task and Rubric Training	Performance Task and Rubric Training
Process 2 (Online)	GenAI and ChatGPT Training	GenAI and ChatGPT ve DALL-E Training	GenAI, ChatGPT and Fliki.ai Training
Post-test	General Attitude towards AI	General Attitude towards AI	General Attitude towards AI
	GenAI Acceptance Scale	GenAI Acceptance Scale	GenAI Acceptance Scale

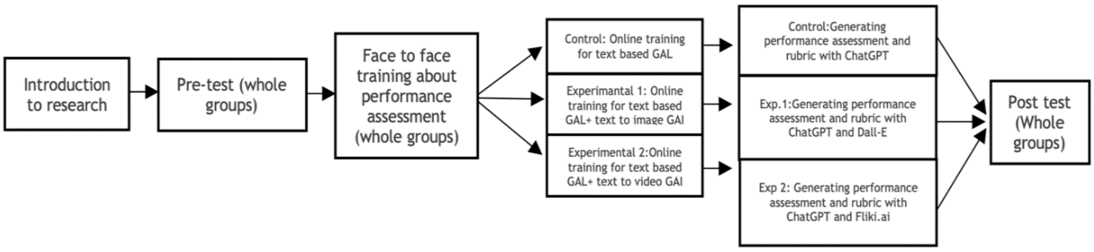

Figure 1**.** Study Process

After the training and pre-test phases, post-tests were administered to each group and the data obtained were collected for analysis (see Figure 1).

GenAI Tools Used in the Study

An analysis of GenAI tools utilized in education reveals the existence of various types, each serving distinct purposes. In addition to text-based GenAI tools such as ChatGPT, Claude, and Gemini, visual content generation systems like MidJourney and DALL-E are playing a significant role in enhancing educational practices. Furthermore, tools that generate multimedia content, including Fliki.ai, Flux.ai, and Sora, enable educators to enrich instructional materials and assessment tools. By leveraging these technologies, educators can offer students more engaging and interactive learning experiences. Research underscores the effectiveness, time-saving capabilities, and user-friendliness of visual AI tools in educational settings (Hisan & Amri, 2023; Vigna-Taglianti, 2024). Although existing studies primarily focus on applications in medical education, they highlight the broader advantages of visual GenAI tools. These tools not only generate royalty-free images of tangible objects but also aid in visualizing complex abstract concepts, thereby facilitating student comprehension (Aktay, 2022). This capability is particularly beneficial in subjects requiring abstract thinking, as it significantly supports the learning process by bridging the gap between conceptual understanding and practical application.

Data Analyses

The research data were analyzed using Jamovi 2.3.28 software. The analysis process was carried out in the following two stages:

Analysis of Pre-Tests

Descriptive statistics (mean, median, mode) were calculated to check the initial similarity of the control and experimental groups. To determine whether the pre-test scores were normally distributed, skewness and kurtosis values were analyzed and the Shapiro-Wilk test was applied. For a data set to be normally distributed, the skewness and kurtosis values should be between -1 and +1 and the Shapiro-Wilk test should not be statistically significant (Field, 2009).

With the exception of the negative attitude sub-dimension of the general attitude towards AI scale, the positive attitude dimension and the acceptance levels of GenAI do not show a normal distribution. Therefore, non-parametric tests such as the Kruskal-Wallis test were used for these variables. In the negative attitude sub-dimension, since it was found that there was a normal distribution and that the group variances were homogeneous, the ANOVA test was preferred for this variable. According to the pre-test results, positive and negative attitudes towards AI were similar in all groups before the experiment. However, significant differences were observed in the levels of acceptance of GenAI. These results were taken into account when analyzing the post-test data.

Analysis of Post-Tests

Descriptive statistics were initially computed in the analysis of the post-test data. The normality of the post-test scores was assessed by analyzing the skewness and kurtosis values, together with the outcomes of the Shapiro-Wilk test. The findings of this normality test indicate that the acceptance ratings for GenAI are not normally distributed ($p < .05$), although the positive and negative attitude scores of AI are regularly distributed ($p > .05$).
Given these normality tests and the analyses of differences between the groups on the pre-test scores, Kruskal-Wallis, ANOVA and ANCOVA tests were used to determine whether the post-test scores differed between the groups. The dependent samples t-test was used for normally distributed data and the Wilcoxon signed-rank test was used for non-normally distributed data to determine the differences between participants' pre- and post-experimental data.

Reliability and Validity

In this study, the reliability of the pre- and post-test scores of the experimental and control groups were assessed using Cronbach's alpha and McDonald's omega coefficients. Table 3 shows the reliability coefficients of the pre- and post-test scores of the control group.

Table 3. Reliability Coefficients for the Pretest and Posttest Scores of the Control Group

	Pre-test		Post-test	
	Cronbach alpha	McDonald's omega	Cronbach alpha	McDonald's omega
Positive attitude towards AI	.911	.916	.854	.871
Negative attitude towards AI	.773	.792	.741	.755
Acceptance of GenAI	.939	.947	.837	.870

The reliability coefficients in Table 3 show values of .70 and above for all variables. These results indicate that the scales have high reliability and the results are reliable (Field, 2009). In Table 4, the reliability coefficients for the pre-test and post-test scores of the Experiment 1 are presented.

Table 4. Reliability Coefficients for Pretest and Posttest Scores for Experiment 1 Group

	Pre-test		Post-test	
	Cronbach alpha	McDonald's omega	Cronbach alpha	McDonald's omega
Positive attitude towards AI	.959	.960	.85	.876
Negative attitude towards AI	.852	.857	.826	.833
Acceptance of GenAI	.987	.984	.951	.955

The reliability coefficients in Table 4 show that the scale scores of the Experiment 1 are highly reliable (Field, 2009). In Table 5, the reliability coefficients for the pre-test and post-test scores of the Experiment 2 are presented.

Table 5. Reliability Coefficients for Pre-test and Post-test Scores for Experiment 2 Group

	Pre-test		Post-test	
	Cronbach alpha	McDonald's omega	Cronbach alpha	McDonald's omega
Positive attitude	.844	.862	.777	.798
Negative attitude	.804	.809	.771	.787
Acceptance of GenAI	.907	.917	.925	.930

Cronbach's alpha and McDonald's omega values in Table 5 show that both the pre-test and post-test scores of the Experiment 2 group are reliable (Field, 2009). The validity of the study is in accordance with the validity standards accepted in the literature for the scales used. The general attitude of AI scale and the GenAI acceptance scale were evaluated in detail in the relevant literature and adapted into Turkish (Kaya et al., 2022; Karacaoglan-Yılmaz et al., 2023). The validity of these scales was assessed in terms of content validity, construct validity and criterion validity.

Ethics

In order to ensure the ethical appropriateness of the research, approval was obtained from the Ethics Committee of the university. The ethics committee approval number is indicated as E-64075176-050.04-240117688. This approval confirms that the research was conducted in accordance with scientific and ethical standards and that participant rights were protected. In addition, participants were asked to sign an informed consent form before participating in the study. This form provided participants with detailed information about the purpose, methods and possible risks of the study.

Results

Descriptive Statistics of Pre-test Scores

Firstly, the attitude towards AI and acceptance levels of GenAI of the control and experimental groups were measured with the descriptive statistics (see Table 6).

Tablo 6. Descriptive Statistics for Pretest Scores for Experimental and Control Groups

Variables	Groups	n	Mean	Median	Standard Deviation	Min	Max
Positive attitude	Control	28	42.2	43.5	7.83	16	56
	Experiment 1	36	43.9	46	10.8	12	60
	Experiment 2	30	45.3	47	7.71	27	59
Negative attitude	Control	28	23	22	5.04	14	38
	Experiment 1	36	22.9	22.5	5.92	9	37
	Experiment 2	30	26	26	6.09	10	39
Acceptance of GenAI	Control	28	71.6	75	11.7	24	88
	Experiment 1	36	73.8	77.5	18.4	21	100
	Experiment 2	30	80.2	77.5	9.30	65	100

As seen in Table 6, Experiment 2 has the highest mean (Mean = 45.3, SD = 7.71). This was followed by the Experiment 1 group (Mean = 43.9, SD = 10.8) and the control group (Mean = 42.2, SD = 7.83). These results show that the positive attitudes of the groups at the beginning were above the middle level. In terms of negative attitude scores, the Experiment 2 had the highest mean (Mean = 26, SD = 6.09). The control group (Mean = 23, SD = 5.04) and Experiment 1 (Mean = 22.9, SD = 5.92) exhibited very close averages. These data show that the negative attitudes of the groups are below the medium level.

Analysis of acceptance levels reveals that the Experiment 2 group has the greatest mean (Mean = 80.2, SD = 9.30). The Experiment 1 has a mean of 73.8 and a standard deviation of 18.4, while the control group has a mean of 71.6 and a standard deviation of 11.7. The results indicate that acceptance levels were elevated across all groups, with the Experimental 2 exhibiting the highest levels compared to the others. Differences were noted among the pretest scores of the three groups. Analyses of differences were conducted to ascertain the statistical significance of these variations, with the results presented in Table 7 and Table 8.

Table 7. Kruskal Wallis Test Results of Pre-test Scores

Variables	Chi-square	df	p
Positive attitude	3.89	2	.143
Acceptance of GenAI	6.76	2	.034*

* $p < 0.05$

Tablo 8. ANOVA Test Results of Pre-test Scores

Variables	F	df1	df2	p
Negative attitude	2.74	2	59.7	.073

As seen in Table 7 and Table 8, it is seen that there is no significant difference between the groups in positive and negative attitude of AI scores. However, a significant difference was found between the groups in the total scores of the acceptance scale for GenAI ($\chi^2 = 6.76$, $p = .034$). The DSCF post-hoc test conducted to determine between which groups this difference was between revealed that there was a significant difference between the control group and the Experiment 2.

Descriptive Statistics of Post-test Scores

After the experimental procedures, post-test was applied to all groups. Descriptive statistics of the post-test scores are presented in Table 9.

Tablo 9. Descriptive Statistics for Post-test Scores for Experimental and Control Groups

Variables	Groups	n	Mean	Median	Std. Deviation	Min	Max
Positive attitude	Control	28	45.9	45.0	5.84	34	60
	Experiment 1	36	48.1	48.0	5.64	34	59
	Experiment 2	30	44.5	44.5	6.52	29	60
Negative attitude	Control	28	22.5	21.5	4.73	12	33
	Experiment 1	36	21.2	20.5	5.41	8	32
	Experiment 2	30	23.5	23.0	5.45	11	40
Acceptance of GenAI	Control	28	78.0	77.5	6.58	67	95
	Experiment 1	36	79.3	78.5	12.0	35	100
	Experiment 2	30	80.8	80.0	10.2	58	100

It is seen that the positive attitude towards AI scores have the highest mean in the Experiment 1 group (Mean = 48.1, SD = 5.64). The control group (Mean = 45.9, SD = 5.84) and Experiment 2 group (Mean = 44.5, SD = 6.52) exhibited similar averages. These data show that the intervention applied in the Experiment 1 group was more effective in increasing the positive attitude towards AI. In order to make sure of this effect, difference analyses between groups were performed (see Table 10).

In terms of negative attitude towards AI scores, small differences were observed between the groups. The

Experiment 2 group had the highest mean (Mean = 23.5, SD = 5.45), followed by the control group (Mean = 22.5, SD = 4.73) and the Experiment 1 group (Mean = 21.2, SD = 5.41). These results suggest that the experimental interventions may have shown a limited effect in reducing negative attitudes. When the acceptance levels of GenAI are analyzed, it is seen that all groups exhibited high averages. Experiment 2 group had the highest mean (Mean = 80.8, SD = 10.2), followed by Experiment 1 group (Mean = 79.3, SD = 12.0) and control group (Mean = 78.0, SD = 6.58). These results suggest that the levels of acceptance of GenAI were high in all groups and that the experimental interventions may have further increased these levels. However, the statistical significance of these differences needs to be examined with further analyses. Therefore, the difference between the control and experimental groups was revealed by Kruskal Wallis and ANOVA tests. ANOVA test was used to evaluate the differences between the post-test scores obtained from the sub-dimensions of the attitude towards AI scale. The results of the analysis are presented in Table 10.

Table 10. ANOVA Results of Post-test Positive and Negative Attitude Scores between Groups

Variables	*F*	*df1*	*df2*	*p*
Positive attitude	2.85	2	58.3	0.066
Negative attitude	1.46	2	59.7	0.241

As seen in Table 10, the ANOVA results of the positive attitude scores show that there is no statistically significant difference between the groups (F(2, 58.3) = 2.85, p = 0.066). However, the fact that the p-value is quite close to 0.05 indicates a significant marginal difference. This indicates that there may be a tendency between the groups in terms of positive attitude, but this tendency does not reach the level of statistical significance.

The ANOVA test for the negative attitude scores revealed that there was no statistically significant difference between the groups (F(2, 59.7) = 1.46, p = 0.241). This result shows that the experimental procedures applied did not have a differentiating effect on the participants' negative attitudes. Nonparametric ANCOVA was employed to analyze the disparities among the groups in the posttest scores for the acceptance levels of GenAI. The findings are presented in Table 11.

Table 11. Non-parametric ANCOVA Results of Post-test Acceptance of GenAI Total Scores between Groups

	Groups (Mean Rank)			Kruskal Wallis H	*df*	*p*
	Control	Experiment 1	Experiment 2			
Acceptance of GenAI	43.2	51.55	47.47	1.358	2	.507

The data in Table 11 indicate that the posttest acceptability levels of participants in both the control and experimental groups were not statistically different (H = 1.358, df = 2, p > .05). Thus, it may be stated that there is no substantial difference between the groups regarding posttest acceptance scores. The various experimental methodologies employed did not appreciably alter the participants' acceptance levels of GenAI. Analysis of the Mean Rank values reveals that the Experiment 1 group possesses the highest value (51.55), succeeded by Experiment 2 (47.47) and the Control group (43.2), respectively. Nonetheless, these disparities lack statistical significance. The p-value of 0.507 suggests that the observed differences between the groups are probably attributable to random variation.

Analyses of the Differences between the Pre-test and Post-test Scores of the Groups

Non parametric Wilcoxon test was preferred for the difference analyses between the pretest-posttest scores of the control group's positive attitude of AI and acceptance levels of GenAI. The Wilcoxon test findings in Table 12 indicate a significant difference between pretest and post-test scores for positive attitudes of AI and acceptance levels of GenAI ($p < .05$). The findings indicate that the posttest scores surpass the pretest scores, and these discrepancies are statistically significant. It was noted that positive opinions increased following the experiment involving the group utilizing solely ChatGPT to complete a performance challenge.Likewise, the acceptability levels of pre-service instructors utilizing solely text-based GenAI tools for performance task creation improved following the experiment.The effect size for the variable of positive attitude was determined to be Cohen's d = -0.59 and rank biserial correlation = -0.71. This signifies a medium to high effect magnitude. The effect size for the acceptability of GenAI was determined as Cohen's d = -0.48 and rank biserial correlation = -0.60. This signifies a moderate effect magnitude. The negative effect size signifies that the post-test results exceed the pretest values.

Table 12. Pre-test-Post-test Difference Analyses of the Control Group's Positive Attitude of AI and Acceptance of GenAI

	Wilcoxon W	*p*	Effect size
Positive attitude	47	.002*	Cohen’s d -0.59
			Rank biserial -0.71
Total acceptance	81.5	.006*	Cohen’s d -0.48
			Rank biserial -0.60

The disparity between the pretest and post-test negative attitude levels in the control group was examined using a dependent samples t-test (see Table 13). It was found that there was no significant difference between the pre-test and post-test scores of the control group's negative attitudes of AI ($t(27)=.484$, $p> .05$). This result shows that there is no change in the negative attitudes of the pre-service teachers who used only ChatGPT while creating the performance task after the experiment. Similar to the control group, pre-test and post-test difference analyses were performed for variables in Experiment 1 group. Wilcoxon test was used for positive attitude of AI and acceptance levels of GenAI.

Table 13. Pre-test-post-test Difference Analyses of the Control Group's Negative Attitude Levels of AI

	Pair sample t test	df	p
Negative attitude	.484	27	.632

No significant difference was found between the pre-test and post-test acceptance levels of the individuals in Experiment 1 group (see Table 14). However, a significant increase was found in the positive attitude levels of the participants in Experiment 1 group ($W= 129$, $p < .05$). In other words, the participants in the group who used both text-based and image-generating AI tool while creating the performance task increased their positive attitudes after the experiment. The effect size was calculated as -0.513 with rank biserial. The negative effect size indicates that the post-test scores were higher than the pretest scores.

Table 14. Pre-test and Post-test Difference Analyses of Positive Attitude and Total Acceptance of Experiment 1 Group

	Wilcoxon W	*p*	Effect size
Positive attitude	129	.011	Rank biserial -0.513
Total acceptance	241	.338	-

The analyses of the level of negative attitude of AI were performed with a dependent samples t-test and given in Table 15.

Table 15. Pre-test and Post-test Difference Analyses of Negative Attitude of Experiment 1 Group

	Pair sample t test	df	p
Negative attitude	1.79	35	.082

In Experiment 1 group, no significant difference was found in negative attitude levels ($p > .05$). In other words, no difference was detected in the attitudes of the participants who used ChatGPT and DALL e while preparing the performance task before and after the experiment. After Experiment 1, the difference between the pre-test and post-test data of the pre-service teachers who used both ChatGPT and fliki.ai platform while creating the performance task was analyzed. The results of these analyses are presented in Table 16 and Table 17.

Table 16. Pre-test and Post-test Difference Analyses of Positive Attitude and Total Acceptance of Experiment 2 Group

	Wilcoxon W	p	Effect size
Positive attitude	243	.367	-
Total acceptance	146.5	.469	-

In the Experiment 2 group, there was no significant difference between the pretest and posttest scores in positive attitude and acceptance levels of GenAI ($p > .05$). This situation indicates that the attitudes and acceptance of the participant group, which used both text-based and video-generating AI tools while creating a performance task, did not differ after the experiment. The dependent samples t-test was applied to compare the pretest and posttest values of the negative attitude levels of the Experimental 2 group (see Table 17).

Table 17. Pre-test-post-test Difference Analyses of the Negative Attitude Levels of the Experiment 2 Group

	Pre-test M	Post-test M	Pair sample t-test	df	p	Effect size (Cohen's d)
Negative attitude	26	23.5	2.497	29	.018	.46

In the Experiment 2 group, a significant difference was found between pre-test and post-test scores in negative attitude levels ($t(29) = 2.497$, $p < .05$). These findings show that the attitudes of those who used both ChatGPT and fliki.ai while creating a performance task changed. When the arithmetic averages of the pre-test and post-test values were analyzed (see Table 17), a decrease was found in the post-test values. This indicates that the negative attitude levels of the participants in Experiment 2 decreased. The Cohen's d value for the effect size was calculated as .46, and according to the classification proposed by Cohen (1988), this value indicates a moderate effect. This result shows that the practical significance of the observed change is also noteworthy.

Discussion

This study examined the impacts of integrating GenAI tools into performance-based task development on pre-service teachers' attitudes and acceptance of AI. The results provide significant insights into the emerging domain of AI in education, especially for the training of future educators. The study revealed high average acceptance levels of GenAI across all groups, both before and after the empirical interventions. This suggests a pre-existing receptiveness among pre-service teachers towards incorporating such tools into their future teaching practices. These findings align with the growing body of research indicating educators' increasing interest in GenAI applications within educational contexts (Darayseh, 2023; Nyaaba et al., 2024; Zhai, 2024). For instance, studies by Darayseh (2023) have also reported high levels of acceptance and positive attitudes towards AI among pre-service teachers. Chung (2024) found that pre-service teachers perceived AI as a valuable tool for enhancing student engagement and personalizing learning. Moreover, Brandhofer and Tengler (2024) demonstrated that familiarity with AI technologies was positively correlated with acceptance levels among educators. The observed increase in positive attitudes, particularly within the control group exposed only to ChatGPT, suggests that hands-on experience with text-generating AI tools may be particularly effective in fostering a more informed and constructive view of their potential benefits in education. This could be attributed to the direct and immediate utility of ChatGPT in generating text-based content relevant to performance task development. However, it is also possible that pre-existing biases or expectations regarding text-based AI tools, compared to more novel visual or video-based tools, contributed to this difference. Future research could explore these potential mediating factors.

Differential Impact of AI Tools and the Technology Acceptance Model (TAM)

Although the post-test results did not reveal statistically significant differences in overall attitudes and acceptance levels between the groups, the analysis of intra-group changes between pre-test and post-test scores presents a more intricate perspective. These results suggest that different types of GenAI tools may not exert homogenous effects on pre-service teachers' perceptions. This discovery emphasizes the intricate nature of AI integration and reveals the possibility for varying effects of different AI tools, supporting Bower et al.'s (2024) conclusions regarding the diversity of AI perceptions among educators influenced by experience, gender, and discipline. The lack of substantial overall group differences underscores the necessity of accounting for individual variances and contextual elements in forthcoming studies, along with Chung's (2024) advocacy for a comprehensive strategy to AI integration.

Delving deeper into the intra-group changes, the control group (using only ChatGPT) exhibited a significant increase in both positive attitudes towards AI and acceptance of GenAI. This demonstrates that even text-based

AI tools like ChatGPT can positively influence pre-service teachers' perceptions of AI's potential benefits and their willingness to use such technologies for pedagogical purposes. The significant increase in positive attitudes towards AI and acceptance of GenAI within the control group resonates strongly with the Technology Acceptance Model (TAM) (Davis, 1989), particularly the construction of perceived usefulness. The experience of using ChatGPT to generate text-based content for performance tasks likely enhanced pre-service teachers' perceptions of its utility in an educational context, thus directly contributing to their behavioral intention to use GenAI in the future, as predicted by TAM (Hamzah et al., 2024; Nair & Das, 2012). Furthermore, the differential impact observed between the groups suggests that different AI tools may influence different constructs within TAM. For example, while ChatGPT may primarily impact perceived usefulness, DALL-E may influence perceived ease of use or even subjective norms, warranting further investigation into these nuanced relationships.

Conversely, the Experiment 1 group, which utilized both text-generating (ChatGPT) and image-generating (DALL-E) AI tools, exhibited a significant increase solely in positive attitudes towards AI. This finding suggests that the integration of image-generating AI tools like DALL-E, while potentially reinforcing the belief in AI's general utility, may not necessarily translate to an increased acceptance of GenAI for direct pedagogical applications at the same rate as text-based tools. This difference could be attributed to the perceived relevance and applicability of the specific AI tool to the task at hand. While ChatGPT directly supports text generation for performance task development, the immediate pedagogical utility of DALL-E might be less apparent to pre-service teachers. This aligns with findings from Yadav (2020) who compared the impact of different AI tools on student learning outcomes and found that the effectiveness of the tool depended on its alignment with the specific learning objectives. Further research by (Hutson & Cotroneo, 2023; Montero, 2024; Wangdi, 2023) has explored the use of visual AI tools in education, highlighting their potential for enhancing creativity and engagement, but also emphasizing the need for careful pedagogical integration (Zebua, 2024). These findings warrant further investigation into the specific mechanisms through which different functionalities of GenAI tools shape teachers' perceptions and acceptance.

A notable outcome was the substantial reduction in negative perceptions of AI among the Experiment 2 group, which utilized both text and video-based AI tools (ChatGPT and Fliki.ai). This indicates that engaging with a tool like Fliki.ai, specifically designed for video creation, might have provided a concrete example of AI's creative potential, thereby mitigating some of the anxieties associated with its perceived negative impacts on critical thinking (Chung, 2024; Echave et al., 2024). This finding aligns with the work of Kim and Kim (2024), who emphasize the importance of effective AI training programs in addressing concerns and fostering informed use of AI tools. The reduction in negative attitudes also suggests that direct and constructive interaction with specific AI tools can play a crucial role in shaping perceptions and alleviating anxieties. This is further supported by research from (Kaur et al., 2024; Matias & Zipitria, 2023) who found that exposure to positive examples of AI applications in education can effectively reduce negative biases. Schiavo et al.'s (2024) study on the impact of AI anxiety on technology acceptance provides further evidence for the importance of addressing negative perceptions to promote successful AI integration.

The intricate results about the varying effects of different AI tools align with the expanding study on the

Technology Acceptance Model (TAM) and its extensions. The original Technology Acceptance Model (Davis, 1989) identifies perceived utility and ease of use as key determinants of technology acceptance, although subsequent models, including Teo's (2009), integrate additional aspects such as self-efficacy and facilitating situations. The findings of this study indicate that various AI tools may have disparate effects on these aspects. The favorable influence of ChatGPT on attitudes and acceptance within the control group may be ascribed to its perceived utility in producing text-based content, consistent with research that has effectively utilized the Technology Acceptance Model (TAM) to forecast technology adoption in educational contexts (Nair & Das, 2012; Hamzah et al., 2024).

Limitations and Future Directions

This study recognizes several limitations, including comparatively small sample size and the reliance on self-reported measures. The limited sample size, sourced from a specific cohort of pre-service teachers, may constrain the applicability of the findings to other groups with differing degrees of AI familiarity and technological proficiency. Subsequent research ought to utilize larger, more heterogeneous samples and integrate observational methodologies to evaluate the influence of AI tool utilization on genuine instructional practices and student outcomes. Examining the enduring impacts of GenAI integration on pre-service teachers' attitudes, acceptance, and pedagogical practices beyond the initial intervention phase is essential. Furthermore, examining the relationship between personal attributes (e.g., previous AI experience, technological self-efficacy, pedagogical views) and the distinct characteristics of various AI tools is crucial for a thorough comprehension of successful AI integration in education.

The findings endorse the integration of GenAI tools into pre-service teacher education curricula. By providing hands-on experiences with these technologies specifically aimed at enhancing performance-based task development, teacher preparation programs can equip future educators with the skills and confidence needed to leverage GenAI effectively into classes. This can ultimately contribute to fostering more engaging and effective learning environments for prospective teachers.

Conclusion

In conclusion, this study demonstrates that the use of GenAI tools in performance task development has complex and varied effects on pre-service teachers' attitudes towards and acceptance of AI. Despite the absence of overall group differences, interaction with different AI tools led to significant changes in intra-group attitude and acceptance levels. The findings of this study underscore the critical need for teacher education programs to adopt a nuanced and differentiated approach to AI integration, one that acknowledges the diverse affordances and limitations of specific AI tools and their potential to impact pre-service teachers' attitudes and acceptance in distinct ways. By meticulously developing interventions that address specific constructs within established theoretical frameworks such as TAM and TPB, teacher educators can proficiently equip future teachers to manage the intricacies of AI in education and leverage its transformative potential to improve teaching and learning.

Statements

Competing Interests: The authors and this candidate article have no competing interests.

Availability of Data and Material: Data stored after the analysis process. It can be shared upon request.

Funding: No financial support was received from any organization or individual during this research process.

References

Ajzen, I., & Fishbein, M. (2000). Attitudes and the attitude-behavior relation: Reasoned and automatic processes. *European Review of Social Psychology, 11*(1), 1–33. https://doi.org/10.1080/14792779943000116

Aksu-Dünya, B., & Yıldız-Durak, H. (2024). Hi! Tell me how to do it: Examination of undergraduate students' chatbot-integrated course experiences. *Quality & Quantity, 58*(4), 3155–3170. https://doi.org/10.1007/s11135-023-01800-x

Aktay, S. (2022). The usability of images generated by artificial intelligence (AI) in education. *International Technology and Education Journal, 6*(2), 51–62. https://dergipark.org.tr/en/pub/itej/issue/75198/1233537

Arguson, A., Mabborang, M., & Paculanan, R. (2023). The acceptability of generative ai tools of selected senior high school teachers in schools' divisions office-manila, philippines. *Journal of Artificial Intelligence, Machine Learning and Neural Network*, *3*(6), 1-10. 3. https://doi.org/10.55529/jaimlnn.36.1.10

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. https://doi.org/10.1007/s10639-019-09968-2

Bannister, P., Santamaría-Urbieta, A., & Alcalde-Peñalver, E. (2023). A Delphi Study on Generative Artificial Intelligence and English Medium Instruction Assessment: Implications for Social Justice. *Iranian Journal of Language Teaching Research*, *11*(3), 53-80.

Berk, R. A. (1986). *Performance assessment: Methods & applications.* Johns Hopkins University Press.

Bezjak, S. (2024). Perceptions and perspectives: Understanding teachers' attitudes towards AI in education. *Development, 23,* 25. https://toknowpress.net/ISBN/978-961-6914-31-4/68.pdf

Bower, M., Torrington, J., Lai, J. W., Petocz, P., & Alfano, M. (2024). How should we change teaching and assessment in response to increasingly powerful generative artificial intelligence? Outcomes of the ChatGPT teacher survey. *Education and Information Technologies,* 1–37. https://doi.org/10.1007/s10639-023-12405-0

Bozkurt, A. (2023). Generative artificial intelligence (AI) powered conversational educational agents: The inevitable paradigm shift. *Asian Journal of Distance Education*, *18*(1).

Brandhofer, G., & Tengler, K. (2024). Acceptance of artificial intelligence in education: Opportunities, concerns, and need for action. *Advances in Mobile Learning Educational Research, 4*(2), 1105–1113. https://doi.org/10.25082/AMLER.2024.02.005

Bugti, F., Sarhandi, S. A., & Bugti, S. M. (2024). A qualitative study on challenges and opportunities for technology integration in pedagogy. *Journal of Development and Social Sciences, 5*(3), 100–109. https://doi.org/10.47205/jdss.2024(5-III)10

Cabero-Almenara, J., Palacios-Rodríguez, A., Loaiza-Aguirre, M. I., & Rivas-Manzano, M. D. R. D. (2024). Acceptance of educational artificial intelligence by teachers and its relationship with some variables and pedagogical beliefs. *Education Sciences*, *14*(7), 740.

Chandrasekera, T., Hosseini, Z., & Perera, U. (2024). Can artificial intelligence support creativity in early design processes? *International Journal of Architectural Computing,* 0(0). https://doi.org/10.1177/14780771241254637

Chiu, T. K. F. (2023). The impact of generative AI (GenAI) on practices, policies, and research direction in education: A case of ChatGPT and Midjourney. *Interactive Learning Environments,* 1–17. https://doi.org/10.1080/10494820.2023.2253861

Chiu, T. K., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence, 4*, 100–118. https://doi.org/10.1016/j.caeai.2022.100118

Chung, C. J. (2024). Preservice Teachers' Perceptions of AI in Education. *AI-EDU Arxiv*, 1-5. https://doi.org/10.36851/ai-edu.vi0.4155

Cohen, J. (1988). Statistical power analysis for the behavioral sciences, 2nd ed. Erlbaum.

Conrad, E. J., & Rees, L. K. (2024). Generative AI as a catalyst to elevate school health education. *Journal of Physical Education, Recreation & Dance, 95*(6), 45–51. https://doi.org/10.1080/07303084.2024.2355865

Crawford, J., Cowling, M., & Allen, K. A. (2023). Leadership is needed for ethical ChatGPT: Character, assessment, and learning using artificial intelligence (AI). *Journal of University Teaching & Learning Practice, 20*(3), 02.

Creswell, J. W., & Creswell, J. D. (2020). *Research design: Qualitative, quantitative, and mixed methods approach*. Sage publications.

Darayseh, A. (2023). Acceptance of artificial intelligence in teaching science: Science teachers' perspective. *Computers and Education: Artificial Intelligence*, *4*, 100132.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use and user acceptance of information technology. *MIS Quarterly 13*(3), 319–339.

Driessens, O., & Pischetola, M. (2024). Danish university policies on generative AI: Problems, assumptions, and sustainability blind spots. *MedieKultur: Journal of Media and Communication Research, 40*(76), 31–52. https://doi.org/10.7146/mk.v40i76.143595

Echave, J. A., Donato, V. J. P., & Acosta, O. D. (2024). Embracing Generative AI in Education: Exploring Teachers' Perceptions, Practices, and Potential Impact. *Jurnal Pendidikan Indonesia Gemilang*, *4*(2), 211-232.

Evmenova, A. S., Borup, J., & Shin, J. K. (2024). Harnessing the power of generative AI to support ALL learners. *TechTrends,* 1–12. https://doi.org/10.1007/s11528-024-00966-x

Field, A. (2009). Discovering statistics using SPSS. Sage Publications Ltd.

Gunawan, K. D. H., Kaniawati, I., Setiawan, W., Rochintaniawati, D., & Sinaga, P. (2021, March). Profile of teachers' integrated science curricula that support by intelligent tutoring systems. *Journal of Physics:*

Conference Series, 1806(1), 012139. https://doi.org/10.1088/1742-6596/1806/1/012139

Halagatti, M., Gadag, S., Mahantshetti, S., Hiremath, C. V., Tharkude, D., & Banakar, V. (2023). Artificial intelligence: the new tool of disruption in educational performance assessment. In *Smart Analytics, Artificial Intelligence and Sustainable Performance Management in a Global Digitalised Economy* (pp. 261-287). Emerald Publishing Limited.

Hamzah, F., Abdullah, A. H., & Ma, W. (2024). Advancing education through technology integration, innovative pedagogies and emerging trends: A systematic literature review. *Journal of Advanced Research in Applied Sciences and Engineering Technology, 41*(1), 44–63. https://doi.org/10.37934/araset.41.1.4463

Hisan, U. K., & Amri, M. M. (2023). ChatGPT and medical education: A double-edged sword. Journal of Pedagogy and Education Science,2(01), 71–89. https://doi.org/10.56741/jpes.v2i01.302

Hutson, J., & Cotroneo, P. (2023). Generative AI tools in art education: Exploring prompt engineering and iterative processes for enhanced creativity. *Metaverse, 4*(1), 14.

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

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.

Karacaoglan-Yilmaz, F. G., Yilmaz, R., & Ceylan, M. (2023). Generative artificial intelligence acceptance scale: A validity and reliability study. *International Journal of Human–Computer Interaction,* 1–13. https://doi.org/10.1080/10447318.2023.228873

Kaur, S., Budhraja, K., Pahuja, A., Nayyar, V., & Saluja, S. (2024). Leveraging Artificial Intelligence in Education: Enhancing Learning Experience. In *Ethical AI and Data Management Strategies in Marketing* (pp. 125-140). IGI Global. https://doi.org/10.4018/979-8-3693-6660-8.ch010

Kaya, F., Aydin, F., Schepman, A., Rodway, P., Yetişensoy, O., & Demir Kaya, M. (2022). The roles of personality traits, AI anxiety, and demographic factors in attitudes towards artificial intelligence. *International Journal of Human–Computer Interaction.* https://doi.org/10.1080/10447318.2022.2151730

Kim, I., & Kim, K. (2024). A Study on Contents Development for the Use of Generative AI in Elementary and Secondary Classes. *Journal of The Korea Society of Computer and Information*, *29*(8), 223-230. https://doi.org/10.9708/jksci.2024.29.08.223

Kim, S. W. (2023). Change in attitude toward artificial intelligence through experiential learning in artificial intelligence education. *International Journal on Advanced Science, Engineering & Information Technology, 13*(5). https://doi.org/10.1007/s10639-023-12172-y

Labadze, L., Grigolia, M., & Machaidze, L. (2023). Role of AI chatbots in education: Systematic literature review. *International Journal of Educational Technology in Higher Education, 20*(1). https://doi.org/10.1186/s41239-023-00426-1

Latif, E., & Zhai, X. (2024). Fine-tuning ChatGPT for automatic scoring. *Computers and Education: Artificial Intelligence, 6*, 100210. https://doi.org/10.1016/j.caeai.2024.100210

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

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, 17*(24), 144–156. https://doi.org/10.3991/ijet.v17i24.36037

Liu, Q., Zhuang, Y., Bi, H., Huang, Z., Huang, W., Li, J., Yu, J., Liu, Z., Hu, Z., Hong, Y., Pardos, Z. A., Ma, H., Zhu, M., Wang, S., & Chen, E. (2024). Survey of computerized adaptive testing: A machine learning perspective. *arXiv preprint arXiv:2404.00712*. https://doi.org/10.48550/arXiv.2404.00712

Lu, H., He, L., Yu, H., Pan, T., & Fu, K. (2024). A study on teachers' willingness to use generative AI technology and its influencing factors: Based on an integrated model. *Sustainability, 16*(16), 7216. https://doi.org/10.3390/su16167216

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education.* Pearson.

Matias, A., & Zipitria, I. (2023). Promoting Ethical Uses in Artificial Intelligence Applied to Education. In *International Conference on Intelligent Tutoring Systems* (pp. 604-615). Cham: Springer Nature Switzerland.

Metin, M. (2013). Teachers' difficulties in preparation and implementation of performance task. *Educational Sciences: Theory and Practice, 13*(3), 1664-1673. Doi: 10.12738/estp.2013.3.1452

Montero, J. B. (2024). Revolutionizing creativity: Unleashing the power of AI in upper elementary art education. *International Journal of Learning and Teaching, 10*(4), 510-515. https://doi.org/10.18178/ijlt.10.4.510-515

Moorhouse, B. L., Yeo, M. A., & Wan, Y. (2023). Generative AI tools and assessment: Guidelines of the world's top-ranking universities. *Computers and Education Open*, *5*, 100-151. https://doi.org/10.1016/j.caeo.2023.100151

Mujtaba, D. F., & Mahapatra, N. R. (2020, December). Artificial intelligence in computerized adaptive testing. In *2020 International Conference on Computational Science and Computational Intelligence (CSCI)* (pp. 649-654). IEEE.

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., Shi, L., Nabang, M., Zhai, X., Kyeremeh, P., Ayoberd, S. A., & Akanzire, B. N. (2024). Generative AI as a Learning Buddy and Teaching Assistant: Pre-service Teachers' Uses and Attitudes. *arXiv preprint arXiv:2407.11983*.

Ofosu-Ampong, K. (2024). Artificial intelligence research: A review on dominant themes, methods, frameworks and future research directions. *Telematics and Informatics Reports*, 100127.

Ou, A. W., Stöhr, C., & Malmström, H. (2024). Academic communication with AI-powered language tools in higher education: From a post-humanist perspective. *System, 121*, 103225. https://doi.org/10.1016/j.system.2023.103225

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

Palm, T. (2008). Performance assessment and authentic assessment: A conceptual analysis of the literature. *Practical Assessment, Research & Evaluation, 13*(4). https://scholarworks.umass.edu/pare/vol13/iss1/4

Pettersson, J., Hult, E., Eriksson, T., & Adewumi, T. (2024). Generative AI and Teachers--For Us or Against Us? A Case Study. *arXiv preprint arXiv:2404.03486*.

Picton, I., & Clark, C. (2024). Children and Young People's Use of Generative AI to Support Literacy in 2024. *National Literacy Trust*.

Qadir, J. (2023). 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.

Santos, R. P. D. (2023). Enhancing physics learning with ChatGPT, bing chat, and bard as agents-to-think-with: a comparative case study. *arXiv preprint arXiv:2306.00724*.

Schepman, A., & Rodway, P. (2023). The General Attitudes towards Artificial Intelligence Scale (GAAIS): Confirmatory validation and associations with personality, corporate distrust, and general trust. *International Journal of Human–Computer Interaction*, *39*(13), 2724-2741. https://doi.org/10.1080/10447318.2022.2085400

Schiavo, G., Businaro, S., & Zancanaro, M. (2024). Comprehension, apprehension, and acceptance: Understanding the influence of literacy and anxiety on acceptance of artificial intelligence. *Technology in Society, 77,* 102537. https://doi.org/10.1016/j.techsoc.2024.102537

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

Teo, T. (2009). Modelling technology acceptance in education: A study of pre-service teachers. *Computers & Education*, *52*(2), 302-312. https://doi.org/10.1016/j.compedu.2008.08.006

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/Christian-Terwiesch-Chat-GTP.pdf

Vigna-Taglianti, J. (2024). AI-Generated Images as a Teaching Tool in Foreign Language Acquisition. *«Technology and language»(Технологии в инфосфере)*, *5*(3), 85-105. https://doi.org/10.48417/technolang.2024.03.07

Wangdi, P. (2023). Integrating artificial intelligence in education: Trends and opportunities. *International Journal of Research in STEM Education, 6*(2). https://doi.org/10.33830/ijrse.v6i2.1722

Westera, W., Prada, R., Mascarenhas, S., Santos, P. A., Dias, J., Guimarães, M., ... & Ruseti, S. (2020). Artificial intelligence moving serious gaming: Presenting reusable game AI components. *Education and Information Technologies, 25*, 351–380. https://doi.org/10.1007/s10639-019-09968-2

Xia, L. A., Sun, O. Y., & Reis, C. (2024). Hearing Students' Voices: Access, Use Patterns and Attitudes toward Generative AI Technologies among Chinese Students. *ChatGPT and Global Higher Education: Using Artificial Intelligence in Teaching and Learning*, 122.

Yadav, A. (2020). Assessment of Artificial Intelligence virtual Learning Applications & Conventional Classroom Learning Methods. *Journal of emerging technologies and innovative research, 7*(3):1405-1413-1405-1413. https://www.jetir.org/papers/JETIR2003207.pdf

Yancey, K. P., Laflair, G., Verardi, A., & Burstein, J. (2023). Rating short L2 essays on the CEFR scale with

GPT-4. In E. Kochmar, J. Burstein, A. Horbach, R. Laarmann-Quante, N. Madnani, A. Tack, V. Yaneva, Z. Yuan, & T. Zesch (Eds.), *Proceedings of the 18th Workshop on Innovative Use of NLP for Building Educational Applications* (pp. 576–584). Retrieved from https://aclanthology.org/2023.bea-1.49

Yin, M., Ma, B., & Pan, X. (2024). Consumer attitudes towards GENAI advertisements: a multi-experiment comparative study based on engagement. *Journal of Cases on Information Technology (JCIT), 26*(1), 1-35. https://doi.org/10.4018/JCIT.356502

Yusuf, A., Pervin, N., Román-González, M., & Noor, N. M. (2024). Generative AI in education and research: A systematic mapping review. *Review of Education*, *12*(2), e3489.

Zebua, N. (2024). Evaluating the Impact of GenAI in High School Education: A Critical Review. *Polygon: Jurnal Ilmu Komputer dan Ilmu Pengetahuan Alam*, *2*(6), 79-86. https://doi.org/10.62383/polygon.v2i6.296

Zhai, X. (2024). Transforming teachers' roles and agencies in the era of generative AI: Perceptions, acceptance, knowledge, and practices. *Journal of Science Education and Technology.* https://doi.org/10.1007/s10956-024-10174-0