International Journal of Human–Computer Interaction

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/hihc20

Generative Artificial Intelligence Acceptance Scale: A Validity and Reliability Study

Fatma Gizem Karaoglan Yilmaz, Ramazan Yilmaz & Mehmet Ceylan

INTERNATIONAL JOURNAL OF HUMAN–COMPUTER INTERACTION
https://doi.org/10.1080/10447318.2023.2288730

Generative Artificial Intelligence Acceptance Scale: A Validity and Reliability Study

Fatma Gizem Karaoglan Yilmaz[a], Ramazan Yilmaz[a], and Mehmet Ceylan[b]

[a]Department of Computer Technology & Information Systems, Faculty of Science, Bartin University, Bartin, Turkey; [b]Department of Communication Coordination, Bartin University, Bartin, Turkey

ABSTRACT

The purpose of this study is to formulate an acceptance scale grounded in the Unified Theory of Acceptance and Use of Technology (UTAUT) model. The scale is designed to scrutinize students' acceptance of generative artificial intelligence (AI) applications. This tool assesses students' acceptance levels toward generative AI applications. The scale development study was conducted in three phases, encompassing 627 university students from various faculties who have utilized generative AI tools such as ChatGPT during the 2022–2023 academic year. To evaluate the face and content validity of the scale, input was sought from professionals with expertise in the field. The initial sample group ($n = 338$) underwent exploratory factor analysis (EFA) to explore the underlying factors, while the subsequent sample group ($n = 250$) underwent confirmatory factor analysis (CFA) for the verification of factor structure. Later, it was seen that four factors comprising 20 items accounted for 78.349% of total variance due to EFA. CFA results confirmed that structure of the scale, featuring 20 items and four factors (performance expectancy, effort expectancy, facilitating conditions, and social influence), was compatible with the obtained data. Reliability analysis yielded Cronbach's alpha coefficient of 0.97, and the test–retest method demonstrated a reliability coefficient of 0.95. To evaluate the discriminative power of the items, a comparative analysis was conducted between the lower 27% and upper 27% of participants, with subsequent calculation of corrected item-total correlations. The results demonstrate that the generative AI acceptance scale exhibits robust validity and reliability, thus affirming its effectiveness as a robust measurement instrument.

KEYWORDS
Generative artificial intelligent; ChatGPT; students; technology acceptance; UTAUT model

1. Introduction

Artificial intelligence (AI), a rapidly developing technology in recent years, has made significant contributions to many fields. AI pertains to the capability of computer systems or machines to carry out tasks that typically demand human intelligence (Kaplan & Haenlein, 2019). AI technologies are developed to emulate essential aspects of human cognition, such as reasoning, learning, perception, and problem-solving (Lieto et al., 2018; Tegmark, 2018). Through the application of algorithms and statistical models, AI systems analyze and interpret complex datasets, learning from experience and adapting their actions accordingly (Ghahramani, 2015; Mitchell, 2019). One of the most significant benefits of this technology is that it offers solutions to problems in different areas of society. AI is used in many areas of our lives, facilitating and improving people's lives. Education is one of the areas where innovative applications of AI are used.

AI in education is a fast-developing field with great potential to transform learning and teaching processes (Hwang et al., 2020; Yilmaz et al., 2022). AI aims to yield personalized, adaptive, and effective learning experiences for students in education and supports teachers with data-driven insights and automated management tasks (Hwang & Tu, 2021; Raffaghelli et al., 2022). AI in education is developing rapidly, especially in higher education, with the developments in educational data mining and learning analytics (Yilmaz et al., 2022). AI applications provide many advantages, such as making educational processes more effective, efficient, and attractive, increasing the success levels of students with personalized learning applications, and reducing teachers' workload and seeing them as potential facilitators. AI, in educational settings, can be utilized for several purposes, such as analyzing data on student performance and behavior, providing personalized recommendations for learning materials and activities, developing intelligent learning systems, and automating grading and assessment (Chen et al., 2020; Ouyang & Jiao, 2021; Roll & Wylie, 2016). AI incorporates the application of advanced technologies like machine learning, deep learning, natural language processing (NLP), computer vision, expert systems, learning analytics, and robotics to elevate the quality of teaching and learning processes. Generative AI applications are one of the technologies used recently.

CONTACT Fatma Gizem Karaoglan Yilmaz gkaraoglanyilmaz@gmail.com Department of Computer Technology & Information Systems, Faculty of Science, Bartin University, Bartin, Turkey

Today, with the escalating adoption of technology, generative AI applications are more favored in education. However, the integration of generative AI in education is closely linked to student technology acceptance (Baytak, 2023; Prasad Agrawal, 2023). The Unified Theory of Acceptance and Use of Technology (UTAUT) model provides valuable insights into students' attitudes and behavioral intentions toward using generative AI tools in an educational context. For instance, students must find a generative AI application that is easy to use, as a user-friendly interface is anticipated to enhance students' willingness to engage with such tools. The goal of this research is to formulate a measurement instrument that unveils the acceptance status of generative AI applications among students, specifically within the realm of educational applications.

Studies have shown that the UTAUT model is an effective model for evaluating the acceptance of new technologies by users (Raffaghelli et al., 2022; Teng et al., 2022; Ustun et al., 2023). Therefore, development of a generative AI acceptance tool based on UTAUT model may provide important inferences on the evaluation of the acceptance of these technologies by users. A comprehensive analysis of the literature reveals, it is evident that the volume of research on utilization of generative AI in education is steadily growing (Yilmaz & Karaoglan Yilmaz, 2023a, 2023b). But it is seen that the existing research is mostly review articles. However, there is no research based on a theoretical basis, such as UTAUT model and examining students' acceptance of generative AI. With the generative AI acceptance scale put forward as a result of research, students' acceptance of these tools and applications can be examined. The study, in this respect, offers originality and novelty. Gaining insights into the factors influencing students' acceptance status and implementing appropriate measures based on this understanding will yield valuable contributions to researchers, educators, and policymakers.

2. Background

2.1. Generative AI in education

As an AI method benefits especially in areas such as NLP and image processing (computer vision), generative AI is used to generate new data or content and allows a pre-trained model to generate new data making use of sample data (Harshvardhan et al., 2020; McKinsey, 2023; The New York Times, 2023). This type of AI is designed to acquire the ability to generate novel instances by leveraging input data. It can produce creative outputs with authentic characteristics like natural language or imagery (World Economic Forum, 2023). Generative AI technology can be used in many ways. For example, a NLP model can write new articles or stories that look like real human writing. Likewise, an image processing model can produce realistic images or videos. Creative content can be produced in many areas, such as music, movies, games, and authoring, and works of art can be created via generative AI technology (MIT Technology Review, 2022a, 2022b). Researchers draw attention to the many potential benefits of using large language models in areas such as language learning. It is stated that AI has potential benefits especially in providing feedback and individualized learning (Jeon et al., 2023).

It can be said that generative AI has the potential to revolutionize education in many aspects. The discussion that follows outlines the possible applications of generative AI in the field of education.

The main areas of use of generative AI in education areas such as individualized learning and feedback, intelligent teaching practices, content and material development, measurement and evaluation, knowledge and skill development, guidance and consultancy. The scope and diversity of usage can be expanded through the creative input of both teachers and students. *Personalized learning and feedback:* Generative AI can enable students to create customized learning environments tailored to their needs. Customized content, materials, and environments can be generated by analyzing each student's learning needs and individual differences (Baidoo-Anu & Ansah, 2023; Lo, 2023; Ruiz-Rojas et al., 2023). *Intelligent tutoring:* Generative AI can also facilitate the creation of intelligent teaching systems that offer real-time feedback and guidance, assisting students in mastering complex subjects. These systems provide additional support, content, and materials as needed, adapting to the student's progress. Moreover, they can generate exercises and quizzes suitable for the student's proficiency level, delivering instant feedback and promoting comprehensive learning (Ahmed et al., 2023; Li et al., 2023; Yilmaz et al., 2022). *Content and material creation:* Moreover, it plays a pivotal role in creating content and materials by generating educational resources like e-books, videos, audio materials, images, presentations, quizzes, and interactive simulations. This saves educators time and resources and ensures the production of high-quality, engaging content (Cooper, 2023; Küchemann et al., 2023; Li et al., 2023). *Assessment and evaluation:* It may aid in assessment and grading by automating consistent and objective evaluation of assignments and assessments, thereby alleviating the workload of educators (Fergus et al., 2023; Lo, 2023). *Knowledge and skills development:* It can analyze student speech and writing, providing feedback that helps improve their language and writing skills (Kasneci et al., 2023; Shidiq, 2023). *Guidance and counseling services*: It can be used to make career recommendations for students based on their interests or goals by analyzing student performance and offering feedback to teachers and counselors, assisting them in providing tailored guidance. It can even determine students' intelligence levels, generating test results based on their interests and learning styles, enabling more suitable suggestions for students' academic and career aspirations (Akiba & Fraboni, 2023). Generative AI also proves valuable in tracking, evaluating, and creating personal development plans for students. By monitoring student performance, these algorithms can identify challenges students face in educational settings. Accordingly, suggestions can be made to teachers and counselors to enhance students' academic success. By analyzing student data, generative AI can identify students' strengths and weaknesses, facilitating the creation of personalized development plans (Yilmaz & Karaoglan Yilmaz, 2023a).

2.2. Technology acceptance for generative AI

Generative AI applications can change and redefine roles and responsibilities of students in education. Dissemination of these technologies can enable students to access more data, gain deeper insights into their learning processes, and facilitate more personalized learning experiences. How generative AI can be used in education largely depends on students' creativity. Hence, students' acceptance and use of such technology play an essential role. This research aimed to explore the acceptance status of educational generative AI among students. Within the scope of the research, the UTAUT model is based on technology acceptance and use. Figure 1 displays the UTAUT model.

The UTAUT model is a theoretical model examining factors that affect acceptance and use of technology (Venkatesh et al., 2012). This model is especially made use of in business and technology (Tamilmani et al., 2021). In education, it is also important that students and teachers accept and use educational technologies effectively. Research shows that this model is frequently used to understand and promote the use of technology in education (Teng et al., 2022; Ustun et al., 2023). The UTAUT model is used to examine various factors that affect technology acceptance. These factors include performance expectancy, effort expectancy, facilitating conditions, and social influence (Venkatesh et al., 2012). In education, these factors can be used to understand student and teacher behavior toward technology (Almaiah et al., 2019; Altalhi, 2021; Raffaghelli et al., 2022). *Performance expectancy* pertains to what users expect from technology usage and aim at doing with it (Venkatesh et al., 2003, 2012). For current explore, it can be described as extent to which students perceive that utilizing generative AI will lead to enhanced academic performance. *Effort expectancy* is perception of user's effort to use the technology (Venkatesh et al., 2003; Wang & Wang, 2010). In the context of this research, effort expectancy can be construed as the anticipation that students' interaction with generative AI will involve both physical and mental exertion, essentially capturing the ease of use for generative AI. *Facilitating conditions* refer to the factors that enhance the usability of technology (Venkatesh et al. 2003, 2012). Within the framework of current research, facilitating conditions refer to how students perceive the organizational and technical infrastructure in place to support their effective utilization of generative AI in the context of learning. *Social influence* encompasses other people's attitudes and pressures toward using technology (Venkatesh et al., 2003, 2012). In current research, social influence pertains to students' perceptions of the

Figure 1. Displays the UTAUT model.

expectations from significant individuals (e.g., teachers, peers) regarding their adoption of generative AI in their academic work.

The UTAUT model acknowledges that, alongside the factors above, the characteristics of the technology, user demographics, and cultural aspects also influence the intention to accept and utilize technology. The model further emphasizes the interactions and relationships between these factors. With its comprehensive coverage of technology adoption and use factors, the UTAUT model finds applicability across various industries and sectors. It is commonly used to gain insights into users' acceptance and utilization of technology, particularly in technology development and marketing. Considering these attributes, the UTAUT model was adopted to examine students' acceptance of generative AI for educational purposes. The UTAUT model is typically used to investigate factors that influence users' acceptance of certain technologies. However, the UTAUT model presents a draft structure (Strzelecki, 2023). In this research, the items in the factors of UTAUT model were structured by considering educational uses of generative AI tools and applications. Therefore, the research differs from standard UTAUT model applications. Another point where our research differs from standard practices is that the acceptance of generative AI applications by students will be examined. Education has special requirements compared to other sectors, so understanding the acceptance of generative AI tools used in education is important for educational practice. As a result, the model developed within the scope of the research differs from the standard model applications by structuring the items under the factors in the UTAUT model by considering generative AI tools and applications and student use.

Examination of pertinent research reveals that numerous studies have been conducted on the acceptance of AI in education. Within the purview of these studies, it is evident that multiple research initiatives have been implemented, particularly concentrating on surveillance and security (Park & Jones-Jang, 2022), smart products and marketing (Sohn & Kwon, 2020), accounting applications (Vărzaru, 2022), psycho-social factors (Kelly et al., 2023), and health sciences (Sallam et al., 2023). However, considering the great developments in generative AI technology in the last one or two years, it is understood that new research is needed on the educational acceptance of generative AI tools and applications by students. Strzelecki (2023) developed a model that examines the predictors of ChatGPT adoption and use among higher education students. The relationships between the constructs in the model (performance expectancy, effort expectancy, social influence, facilitating conditions, hedonic motivation, habit, personal innovativeness, behavioral intention, and use behavior) were examined within the scope of the research. Yilmaz et al. (2023) examined the ChatGPT perceptions of students in science and math education programs within the framework of technology acceptance model. In these studies, it is seen that the acceptance status of users for ChatGPT is examined. Researchers state that the acceptance status of other generative AI tools should also be examined in future studies. In our research, as different from these studies, we want to develop a measurement tool that can measure the acceptance status of students in Turkish culture for generative AI in general. At this point, our study differs from the existing studies and has a new and original value.

3. Methods

The primary aim of this study is to develop an instrument based on the UTAUT model to determine the generative AI acceptance of students for educational purposes. During the scale development, participants self-report using a Likert-type rating structure (Tezbaşaran, 1997). Following established protocols, a Likert-type scale was designed within this study to measure the extent of students' acceptance levels toward the utilization of generative AI in educational contexts.

3.1. Participants

The aim in this research is developing a measurement tool to identify users' acceptance of generative AI for educational purposes. Consequently, the study participants were selected to include students with prior experience using generative AI applications, such as ChatGPT, for educational purposes. Participants for the study were selected through purposive sampling from a university in Turkey during spring semester of 2022–2023 academic year.

Purposeful sampling is a deliberate method employed by researchers enabling them to concentrate on a specific target to obtain data relevant to the research topic (Creswell, 2012). It is particularly utilized when the research examines a particular subgroup or population of interest. This approach allows researchers to obtain more comprehensive and precise results (Buyukozturk et al., 2018). Compared to random sampling methods, purposeful sampling is more aligned with the research objectives and is thus more likely to yield accurate findings. However, it is crucial for researchers to be more aware and conscious of potential biases during the sampling to minimize them (Balci, 2001).

The present study employed the purposeful sampling approach by targeting university students using generative AI applications such as ChatGPT for educational purposes. The participants were recruited from various departments within the university where the research was conducted. A data collection tool was developed for students with prior experience using generative AI applications. Prior to engaging in the data collection process, explicit informed consent was secured from all participants involved in the study. The questionnaire, consisting of two sections, was designed in electronic format and distributed to the participants via email. The first section comprehensively explained the research objectives and obtained voluntary participation consent. In contrast, the second section was accessible to those who granted their consent, and contained scale items based on the UTAUT model, which was employed within the scope of this study. While determining students to be included in investigation, first, faculty members in the departments of each faculty and college of the university were interviewed. In the interview with the faculty members, it was tried to determine the departments where generative AI tools were used within the scope of the courses. Then, the researchers met with the students in the relevant departments online and gave information about the purpose of the research and data collection tool. Since the research was conducted on university students in Turkey, data collection tool was developed in Turkish. Students were told that they could voluntarily participate in the study. It was stated that students using generative AI tools such as ChatGPT should participate in the study. After the relevant explanations were made, the web-based measurement tool was sent to the students via e-mail and they were asked to answer the data collection tool. Students from the Faculty of Science, Faculty of Engineering, Architecture and Design, Faculty of Letters, Faculty of Education, Faculty of Sports Sciences, Faculty of Health Sciences, Faculty of Economics and Administrative Sciences, Faculty of Forestry, Faculty of Islamic Sciences, School of Foreign Languages, Vocational Schools and Institute of Graduate Studies participated in the study. In the research, it was tried to create diversity by reaching students from different units. In this way, the acceptance status of the students can be revealed more realistically. When the students' responses to the data collection tools were evaluated, it was seen that the most frequently used generative AI applications were ChatGPT, ChatPDF, Bard, DALL-E, Midjourney, DeepL, Bing AI, and the most widely used one was ChatGPT. It is seen that the age range of the students participating in the study varies between 17 and 55. The average age of the students was 22.07. Students state that they use generative AI tools on their smartphones at a rate of 55.3% and on their laptops at a rate of 41.6%. It is seen that there are various tools such as tablet computers in the remaining rate. When the frequency of students' use of generative AI tools is examined, 15.3% stated that they use them every day, 30.3% several times a week, 7.2% once a week, and 29.2% several times a month. The remaining percentage used it less frequently. When the duration of students' use of generative AI tools is analyzed, it is seen that this period varies between 1 and 6 months.

This research involved three distinct groups of participants. The initial group comprised 338 students, including 176 females and 162 males, from which the data were used for the exploratory factor analysis (EFA), while 250 students, 127 females and 123 males, were involved in the second group. The data acquired from the second group underwent confirmatory factor analysis (CFA). Item analyses and computation of Cronbach's alpha reliability coefficients ($N = 588$) were conducted in both groups. It is advised in the literature to conduct separate sample groups for CFA and EFA (Ilhan & Cetin, 2014). Hence, this study employed distinct participant groups to apply CFA and EFA to the acquired data. According to Kline (1994), it is recommended that the sample size in CFA and EFA should be approximately 10 times the number of items. Therefore, the sample size was kept above 200.

Taking into account, the cost considerations related to administering the educational generative AI acceptance scale repeatedly to the same participant group, test–retest reliability was evaluated via data that a distinct and smaller third group provided. This approach allowed for examining the stability and consistency of the scale's measurements over time while minimizing the resources required. The third group consisted of 39 students, comprising 22 females and 17 males. These students were not included in the groups previously subjected to CFA and EFA. The participants were administered the scale for the first time, and subsequently, to assess test–retest reliability, same test was administered to these students on two separate occasions. There was a three-week interval between the administrations, and the data obtained from both instances were compared to evaluate the consistency and stability of the scale's measurements over time. In test–retest reliability studies, it is stated that a balance should be found when choosing the time interval between measurements. Measurements made between very long intervals may not reflect the true reliability of the variable because the performance of the participants may vary over time. On the other hand, measurements taken over very short periods of time may increase the learning effect of the test and misleadingly overestimate reliability. For this reason, it is stated that an average of three weeks is an ideal period (Buyukozturk et al., 2018). Therefore, a three-week period was taken as a basis for test–retest reliability within the scope of the study.

3.2. Scale

The investigators devised generative AI acceptance scale, incorporating four sub-dimensions based on the UTAUT model, which involve "performance expectancy", "effort expectancy", "facilitating conditions", and "social influence". An extensive literature review was conducted to construct the scale, and suitable items were formulated for each factor. Afterwards, the items were rearranged specifically for the generative AI acceptance scale tailored to educational contexts. As a result of the literature review, the draft items that could be included in the four sub-dimensions within the scope of the UTUAT model were reorganized for generative AI and then submitted to expert opinion. In line with the opinions of the experts (three educational technology experts, two language and expression experts, two measurement and evaluation experts), some draft items were removed from the item pool. Some items were reorganized in line with expert opinions. Inappropriate items were removed after the analyses in the scale development stages conducted after the expert evaluation. In development process, the scale comprised seven items for performance expectancy, five items for effort expectancy, three items for facilitating conditions, and five items for social influence. The reason why sub-dimensions of scale contain different numbers of items is due to the items eliminated in the sub-dimensions during the scale development process. The stages of scale development and each process performed in these stages are explained in detail in the following sections of the article. The responses were of five levels, from “strongly agree” to “strongly disagree” in a Likert scale, based on which participants were asked to show to what extent they agreed or disagreed with the provided statements.

3.3. Procedure

The present study comprehensively evaluated the scale’s face, construct, and content validity. To ascertain face validity and content validity, expert opinions were sought from educational technologists and assessment-evaluation experts. Their perspectives and insights were utilized to assess the scale’s clarity, relevance, and comprehensiveness in relation to the study’s intended purpose and target population. Based on their expert input, one item from the dimensions of effort expectancy and facilitating conditions was revised, along with two items from each of the performance expectancy and social influence dimensions. Subsequently, the scale underwent linguistic evaluation by English and Turkish linguistic experts. Following the revisions, a pilot study involving 22 students was conducted to receive feedback on the scale’s administration duration and item clarity. During a meeting, the students convened and shared their perspectives on the intelligibility of the scale’s instructions and items, leading to required adjustments. The students’ time required for scale completion was determined by considering the average duration of the initial and final respondents. After collecting data from this specific group of participants, their responses were segregated from the primary dataset. The 20-item scale underwent meticulous refinement and revision based on the feedback and insights obtained from this group. The resulting refined version of the scale was deemed the final version to be administered to the main participants in the study.

3.4. Data analysis

To evaluate the scale’s psychometric properties, statistical analyses were conducted following its administration to the main participants. First, EFA was carried out for the data that initial group provided. Preceding conducting the EFA analysis, it is essential to ensure that the dataset fulfills the necessary prerequisites, such as the sample size requirement, as recommended in the literature (Cokluk et al., 2012). The participant group of 338 students meets the minimum sample size criterion for EFA. In order to determine the suitability of the dataset for factor analysis, the Kaiser–Meyer–Olkin (KMO) measure and the Bartlett test were employed. Buyukozturk (2010) suggests that for the dataset to be suitable for factor analysis, the Bartlett test should yield a statistically significant result, and the KMO value should be greater than 0.60. These statistical assessments were conducted to ensure that the dataset met the necessary reliable and valid factor analysis criteria.

EFA encompasses various techniques, including weighted least squares and principal component analyses (PCAs) (Tabachnick & Fidell, 2007). These techniques within EFA allow for a comprehensive analysis of the underlying factors and their relationships within the dataset. These diverse methods within EFA enable a comprehensive exploration of underlying factors and their relationships within the analyzed data. These diverse methods allow for comprehensive exploration and extraction of underlying factors within the data. Of these techniques, PCA is widely acknowledged as the most robust method in terms of psychometrics (Stevens, 1996). Similarly, Akbulut (2010) recommends PCA as the preferred method for factorization. In current examination, the factor analysis was performed utilizing PCA method. PCA is a widely used technique in factor analysis that helps identify the underlying factors and their respective contributions to the observed variables. This study aimed to uncover the latent factors and understand their relationships within the measured variables by employing the PCA method.

To ensure the inclusion of items in the scale, the criterion that each factor should have a factor loading greater than 0.30 was taken into account (Pallant, 2005). Additionally, the measured variables’ commonality values (h^2) play a significant role in item retention. Items with low common variance should be eliminated from measurement instrument. In this study, a cut-off value of 0.20, as Tabachnick and Fidell (2007) recommended, was employed as the minimum threshold for common variance. Therefore, the analysis in this study was performed with a cut-off value of 0.20.

CFA was performed to validate measurement model and support EFA results. Furthermore, standardized values and various fit indices (such as IFI, GFI, TLI, NFI, SRMR, CFI, RMSEA, and Chi-square goodness of fit test) were examined as literature suggests (Byrne, 2010; Hu & Bentler, 1999). The inclusion of these fit indices allows for a comprehensive evaluation of how well the model fits the data. In order to meet the sample size criteria for CFA, as advised by Hair et al. (1979) and Kline (1994), a total of 250 student were included in the study. Two methods were employed to ensure reliability of measurements obtained from generative AI acceptance scale: test–retest reliability and Cronbach’s alpha. Test–retest reliability involved administering the scale to the participants on two separate occasions, with a time interval of three weeks, to assess the consistency of their responses over time. Cronbach’s alpha was computed to evaluate internal consistency of scale. Furthermore, calculation of the corrected item-total correlations and comparison of responses of upper 27% and lower 27% of contributors helped examine the discriminative power of items. This analysis helped determine how well each item differentiated between participants with high and low acceptance levels. The statistical analyses, including test–retest reliability, Cronbach’s alpha, construct validity, item analysis, and EFA, were conducted using SPSS 24.0 software (SPSS Inc., Chicago, IL). Additionally, CFA calculations were performed using AMOS 22.0 software package to validate structure of scale further. These software tools facilitated the comprehensive data analysis and provided robust statistical findings.

Table 1. EFA results of generative AI acceptance scale.

Factors		1	2	3	4	H^2	AVE	CR
Performance expectancy	Item 5	0.815				0.82	0.54	0.89
	Item 6	0.764				0.74		
	Item 2	0.719				0.68		
	Item 7	0.717				0.71		
	Item 3	0.711				0.73		
	Item 4	0.704				0.63		
	Item 1	0.686				0.7		
Social influence	Item 17		0.875			0.87	0.74	0.93
	Item 18		0.870			0.89		
	Item 16		0.853			0.84		
	Item 19		0.852			0.86		
	Item 20		0.838			0.82		
Effort expectancy	Item 10			0.859		0.86	0.66	0.91
	Item 8			0.856		0.84		
	Item 9			0.816		0.84		
	Item 11			0.803		0.82		
	Item 12			0.712		0.72		
Facilitating conditions	Item 14				0.851	0.81	0.54	0.77
	Item 15				0.734	0.8		
	Item 13				0.588	0.69		
	% eigenvalue (total = 15.67)	4.765	4.467	4.368	2.070			
	% variance explained (total = 78.349)	23.824	22.335	21.840	10.350			

4. Findings

4.1. Construct validity

4.1.1. EFA

The implementation of EFA involved use of data from initial group. KMO measure was employed to determine sample size requirement, and Bartlett's test of sphericity was subsequently used to assess dataset's suitability for factor analysis. Analysis found that an item exhibited a high factor loading in both the dimensions of social influence (0.558) and facilitating conditions (0.546), leading to its removal. As a result of a further EFA, it was revealed that another item demonstrated a high factor loading in both the social influence (0.490) and facilitating conditions (0.553) dimensions, prompting its elimination and subsequent repetition of the EFA. Following the EFA, the scale items were individually assigned to their respective dimensions, and the items in the scale were renumbered under one dimension. The KMO value for scale was computed as 0.949, and Bartlett's test of sphericity yielded a significant result ($\chi^2(190) = 6302.080$, $p = 0.000$), indicating dataset's suitability for EFA. To elucidate factor pattern of scale, a PCA with varimax rotation was performed. The findings of EFA are given in Table 1.

Table 1 illustrates that all 20 items exhibit factor loadings exceeding 0.30. Moreover, each item in the scale surpasses the criterion of 0.20 for explained common factor variance. The first dimension, performance expectancy, consists of seven items with factor loadings ranging from 0.686 to 0.815, accounting for 23.824% of the total variance. Comprising five items with factor loadings between 0.838 and 0.875, the second dimension, social influence, explains 22.335% of total variance. The third dimension, effort expectancy, explains 21.840% of variance and consists of five items with factor loadings 0.712. In closing, facilitating conditions dimension contributes to 10.350% of variance and includes three items with factor loadings varying from 0.588 to 0.851. Cumulative variance explained by scale amounts to 78.349%, indicating its effectiveness in capturing the measured construct. Then, the EFA yielded a robust four-factor structure comprising a total of 20 items.

Average variance extracted (AVE) is a value that measures the similarity between the items of a factor. Fornell and Larcker (1981) recommended an AVE greater than 0.5. Here, the AVE value of all four factors is sufficient. Composite reliability (CR) is a value that measures the internal consistency of a factor. Fornell and Larcker (1981) recommended a CR value of 0.60 or more. Here, the CR value of all four factors is sufficient.

4.1.2. CFA

The data collected from second group of participants were employed to validate derived 20-item, four-factor structure obtained from the EFA. This validation process aimed to confirm the stability and consistency of the factor structure across different samples. CFA results yielded following fit indices for scale: $\chi^2 = 2.113$, CFI = .97, GFI = .88, IFI = .97, TLI = .97, RMSEA = .067, and SRMR = .0332. Table 2 presents the acceptable and perfect fit values except GFI (Hu & Bentler, 1999). Even though GFI value is not within the acceptable limit values, it remains in close proximity to the acceptable values, and the other indices demonstrate a perfect fit. These findings suggest that the four-factor model derived from the CFA exhibits acceptable goodness of fit.

The factor loadings for model resulting from the CFA are depicted in Figure 2. As depicted in Figure 2, factor loadings range from 0.84 to 0.93 for performance expectancy factor, between 0.88 and 0.95 for social influence factor, between 0.87 and 0.95 for effort expectancy factor, and between 0.84 and 0.89 for factor of facilitating conditions.

4.2. Reliability

Reliability of scale was determined using Cronbach's alpha and test–retest methods. The calculated reliability

Table 2. Values of goodness of fit index.

	Fit indices obtained	Perfect fit	Acceptable fit	References
	2.113	$0 \leq \chi^2/df \leq 3$	$3 < \chi^2/df \leq 5$	Kline (2005) and Sumer (2000)
CFI	0.97	$0.95 \leq CFI \leq 1$	$0.90 \leq CFI < 0.95$	Tabachnick and Fidell (2007)
GFI	0.88	$0.95 \leq GFI \leq 1$	$0.90 \leq GFI < 0.95$	Miles and Shevlin (2007) and Tabachnick and Fidell (2007)
IFI	0.97	$0.95 \leq IFI \leq 1$	$0.90 \leq IFI < 0.95$	Cokluk et al. (2012)
TLI	0.97	$0.95 \leq TLI \leq 1$	$0.90 \leq TLI < 0.95$	Cokluk et al. (2012)
RMSEA	0.067	$.00 \leq RMSEA \leq 0.05$	$0.05 < RMSEA \leq 0.08$	Hooper et al. (2008)
SRMR	0.0332	$.00 \leq SRMR \leq 0.05$	$0.05 < SRMR \leq 0.10$	Cokluk et al. (2012)

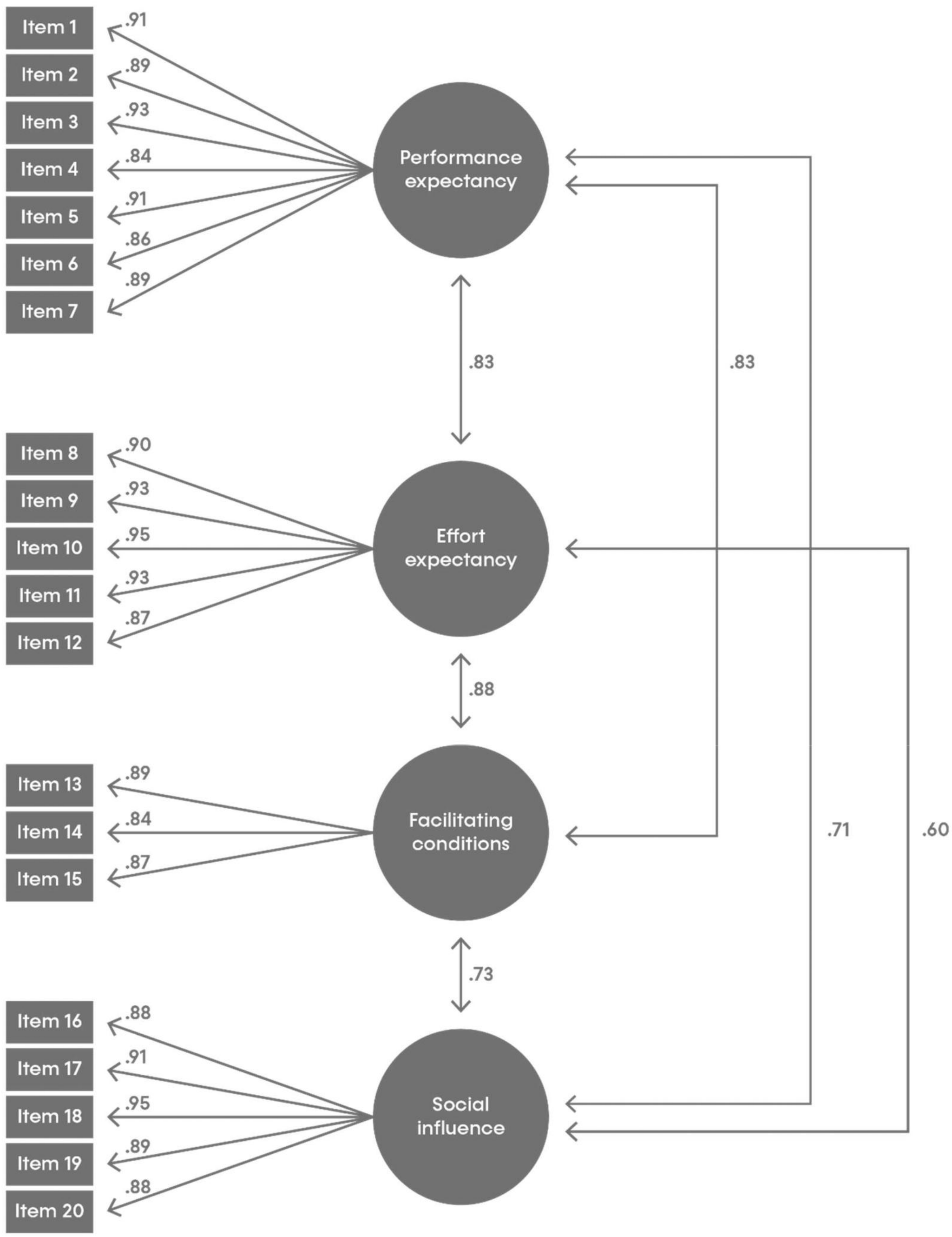

Figure 2. Standardized solutions for generative AI acceptance scale.

Table 3. Reliability coefficients of generative AI acceptance scale.

Sub-dimensions	Cronbach's alpha	Test–retest
Performance expectancy	.96	.93
Social influence	.96	.94
Effort expectancy	.96	.92
Facilitating conditions	.87	.91
Generative AI acceptance scale	.97	.95

coefficients, as shown in Table 3, surpass 0.70 threshold proposed by Fraenkel et al. (2012), indicating the scale's reliability.

When scrutinizing Table 3, it can be observed that Cronbach's alpha and test–retest reliability values calculated for generative AI acceptance scale and dimensions are .87

Table 4. Scale item analysis results.

Item no.	Cronbach's alpha if item deleted	Corrected item-total correlation	Mean	SD	t
Performance expectancy					
Item 1	0.966	0.813	4.0289	.92339	−21.087*
Item 2	0.967	0.791	3.8793	.94159	−21.105*
Item 3	0.966	0.829	4.0918	.90337	−22.237*
Item 4	0.967	0.750	3.8639	.97596	−21.460*
Item 5	0.966	0.821	4.0952	.88395	−22.110*
Item 6	0.967	0.780	3.9830	.93742	−21.706*
Item 7	0.966	0.810	3.9983	.93484	−22.017*
Social influence					
Item 16	0.967	0.750	3.5425	1.06352	−24.304*
Item 17	0.967	0.745	3.5408	1.03105	−23.440*
Item 18	0.967	0.767	3.5816	1.02031	−23.775*
Item 19	0.967	0.725	3.5476	1.06330	−22.842*
Item 20	0.968	0.705	3.5884	1.07820	−22.148*
Effort expectancy					
Item 8	0.967	0.746	4.0204	.95535	−17.804*
Item 9	0.966	0.802	4.0782	.88562	−19.712*
Item 10	0.967	0.782	4.0442	.90319	−19.909*
Item 11	0.967	0.785	4.0238	.89449	−21.162*
Item 12	0.967	0.781	3.9983	.90710	−21.254*
Facilitating conditions					
Item 13	0.967	0.780	3.9507	.91511	−20.763*
Item 14	0.968	0.647	3.8163	.99587	−15.406*
Item 15	0.967	0.763	3.9201	.96906	−19.280*

*$p < 0.01$.

and above. This finding shows that reliability of scale is high.

4.3. Item analysis

The outcomes of the item analysis, featuring adjusted item-total correlations, are detailed in Table 4. The purpose of these correlations was to assess how effectively the scale items could discriminate and predict the overall score. A comparison was conducted between upper 27% and lower 27% of participants to examine differential performance of items across these groups. This analysis helps to determine extent to which individual items contribute to differentiating between participants with high and low scores on the scale.

Table 4 reveals that significant t-values were observed for all items when comparing the upper and lower 27% of students in each dimension. Specifically, for the performance expectancy, the t-values were −21.087 and −22.237, while they were −22.148 and −24.304 for the social influence dimension. The effort expectancy dimension yielded t-values of −17.804 and −21.254, while facilitating conditions showed t-values of −15.406 and −20.763. The comparison between the upper and lower groups for the significant t-values indicates discriminative power of items (Buyukozturk, 2010).

Results presented in Table 4 indicate that item-total correlations vary across dimensions of scale. Specifically, the correlations range from 0.750 to 0.829 for the performance expectancy dimension. For the social influence dimension, the correlations range from 0.705 to 0.767. The effort expectancy dimension shows correlations ranging from 0.746 to 0.802, while the facilitating conditions dimension demonstrates correlations ranging from 0.647 to 0.780. According to the interpretation of item-total correlations, values of 0.30 and above indicate sufficient discriminative power (Buyukozturk, 2010). It is noteworthy that all items in the scale meet this requirement, confirming their discriminative power.

4.4. Correlations between subdimensions

Table 5 presents analysis findings conducted to explore relationship between sub-dimensions of scale.

Table 5 reveals that correlations among sub-dimensions of scale range from 0.550 to 0.768 and are statistically significant at the 0.01 level.

4.5. Interpreting the scores of generative AI acceptance scale

The generative AI acceptance scale is a 20-item one with responses of participants in five levels from "strongly disagree" to "strongly agree" (Appendix 1). The total scores obtained from the scale range from 20 to 100. Higher scores indicate greater acceptance of generative AI among students.

5. Discussion and conclusions

The interest in the utilization of generative AI across diverse domains has witnessed rapid growth, with novel applications being introduced to users on a daily basis. The effective integration of generative AI within the field of education offers substantial opportunities. However, a fundamental prerequisite for harnessing the potential benefits of this technology is the acceptance of generative AI by its users (Sezer & Yilmaz, 2019). Consequently, in order to fully capitalize on the advantages offered by generative AI, students must demonstrate acceptance toward its adoption. The present study aimed to develop a generative AI acceptance scale specifically designed for evaluating students' acceptance of generative AI applications within an educational context.

Table 5. Pearson correlations between sub-dimensions.

		Performance expectancy	Social influence	Effort expectancy	Facilitating conditions	Total scale
Performance expectancy	r	–				
Social influence	r	0.658**	–			
Effort expectancy	r	0.768**	0.550**	–		
Facilitating conditions		0.719**	0.630**	0.741**	–	
Total scale		0.924**	0.829**	0.869**	0.848**	–

**Correlation is significant at the 0.01 level (two-tailed).

The scale was developed based on the established UTAUT model. Extensive literature review indicates that the UTAUT model has been widely employed in numerous studies investigating the acceptance of diverse technologies (Chang et al., 2022; Raffaghelli et al., 2022; Toh et al., 2023). Recognized for its robustness and comprehensive nature, the UTAUT model elucidates acceptance and utilization of technology (Venkatesh et al., 2012). Thus, choice of this model as conceptual framework for this current research aims to bridge the research gap and formulate a generative AI acceptance scale.

The present research adhered to standard procedures and protocols in the development of the measurement scale. Both CFA and EFA were employed to assess the construct validity. EFA yielded a four-factor structure comprising 20 items, explaining 78.349% of total variance. The CFA was employed to assess measurement model's accuracy, and fit indices indicated an acceptable fit for the construct. All items exhibited factor loadings exceeding the cutoff criterion of 0.30, further supporting the construct validity established through CFA. Internal consistency of the scale was assessed using Cronbach's alpha, which yielded a value of 0.97. Additionally, test–retest reliability was evaluated to ascertain the scale's stability, resulting in a coefficient of 0.95. Reliability coefficients of 0.70 or higher are deemed satisfactory according to the literature (Fornell & Larcker, 1981), thus affirming the high reliability of developed instrument. Item analysis, utilizing corrected item-total correlations, was employed to assess the predictive and discriminative power of the scale items. Comparisons were made between upper and lower quintiles of students. This study's findings indicated that all items possessed discriminative power. Consequently, the instrument proves to be a valid and reliable scale for evaluating students' intention to adopt generative AI.

It is imperative to acknowledge the existence of limitations in this study. A systematic and step-by-step approach was utilized to conduct validity and reliability testing of scale, whereas generalizability of findings is subject to some limitations that need to be addressed in future studies. The participants in this research consisted of university students from all class levels at various faculties of a state university. To enhance external validity of results, future investigations could involve participants from different universities. In future research, the acceptance status of different generative AI tools by students can be examined comparatively. This research reflects the results of students at a university in Turkey. Therefore, there is another limitation due to the country-specific nature of the dataset (Jang et al., 2022). Cultural variations may constrain the generalization of the study's outcomes. To reveal and understand cross-cultural differences, future research could incorporate cross-cultural studies. One of the important points in generative AI is related to ethical production and use. Students' acceptance status can be examined in the context of ethical production and use. The effect of individual differences on students' acceptance of generative AI can be examined. Again, in further research, it can be investigated whether the use of generative AI tools will cause privacy concerns in students. In Park's (2021) research that AI can affect users' perceptions of privacy, and that the fear of being monitored can be reflected in individual psychology, institutional behaviors, and policy principles. Since research was ran on university students in Turkey, data collection tool was developed in Turkish and the final version was translated into English. In future studies, the reliability of the English translation of the scale can be examined on students in English-speaking countries. In the study, students' acceptance behaviors were examined in the context of performance expectancy, effort expectancy, facilitating conditions, and social influence factors. In future research, effects of variables such as generative AI tool type, individual differences of users on performance expectancy, effort expectancy, facilitating conditions, and social influence can be investigated.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Fatma Gizem Karaoglan Yilmaz http://orcid.org/0000-0003-4963-8083
Ramazan Yilmaz http://orcid.org/0000-0002-2041-1750

References

Ahmed, F., Shubeck, K., & Hu, X. (2023, July). ChatGPT in the generalized intelligent framework for tutoring. *Proceedings of the 11th Annual Generalized Intelligent Framework for Tutoring (GIFT) Users Symposium (GIFTSym11)* (p. 109). US Army Combat Capabilities Development Command-Soldier Center.

Akbulut, Y. (2010). *Sosyal bilimlerde SPSS uygulamaları [SPSS applications in social sciences]*. İdeal Kültür Yayıncılık.

Akiba, D., & Fraboni, M. C. (2023). AI-supported academic advising: Exploring ChatGPT's current state and future potential toward student empowerment. *Education Sciences*, *13*(9), 885. https://doi.org/10.3390/educsci13090885

Almaiah, M. A., Alamri, M. M., & Al-Rahmi, W. (2019). Applying the UTAUT model to explain the students' acceptance of mobile learning system in higher education. *IEEE Access*, *7*, 174673–174686. https://doi.org/10.1109/ACCESS.2019.2957206

Altalhi, M. (2021). Toward a model for acceptance of MOOCs in higher education: The modified UTAUT model for Saudi Arabia.

Education and Information Technologies, *26*(2), 1589–1605. https://doi.org/10.1007/s10639-020-10317-x

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.

Balci, A. (2001). *Sosyal bilimlerde araştırma. Yöntem, teknik ve ilkeler [Research in the social sciences. Methods, techniques and principles].* PegemA Yayınları.

Baytak, A. (2023). The acceptance and diffusion of generative artificial intelligence in education: A literature review. *Current Perspectives in Educational Research*, *6*(1), 7–18.

Buyukozturk, S. (2010). *Sosyal bilimler için veri analizi el kitabı [The data analysis handbook for social sciences].* Pegem Akademi Yayınları.

Buyukozturk, S., Kilic Cakmak, E., Akgun, O. A., Karadeniz, S., & Ve Demirel, F. (2018). *Bilimsel araştırma yöntemleri [Scientific research methods].* Pegem.

Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming.* Taylor & Francis.

Chang, M., Walimuni, A. C., Kim, M. C., & Lim, H. S. (2022). Acceptance of tourism blockchain based on UTAUT and connectivism theory. *Technology in Society*, *71*(2022), 102027. https://doi.org/10.1016/j.techsoc.2022.102027

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

Cokluk, O., Sekercioglu, G., & Buyukozturk, S. (2012). *Sosyal bilimler için çok değişkenli istatistik: SPSS ve LISREL uygulamaları [Multivariate statistics for the social sciences: SPSS and LISREL applications].* Pegem Akademi Yayıncılık.

Cooper, G. (2023). Examining science education in ChatGPT: An exploratory study of generative artificial intelligence. *Journal of Science Education and Technology*, *32*(3), 444–452. https://doi.org/10.1007/s10956-023-10039-y

Creswell, J. W. (2012). *Educational research: Planning, conducting and evaluating quantitative and qualitative research.* Pearson.

Fergus, S., Botha, M., & Ostovar, M. (2023). Evaluating academic answers generated using ChatGPT. *Journal of Chemical Education*, *100*(4), 1672–1675. https://doi.org/10.1021/acs.jchemed.3c00087

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50. https://doi.org/10.1177/002224378101800104

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education* (8th ed.). McGraw-Hill.

Ghahramani, Z. (2015). Probabilistic machine learning and artificial intelligence. *Nature*, *521*(7553), 452–459. https://doi.org/10.1038/nature14541

Hair, J. F., Anderson, R. E., Tatham, R. L., & Grablowsky, B. J. (1979). *Multivariate data analysis*. Pipe Books.

Harshvardhan, G. M., Gourisaria, M. K., Pandey, M., & Rautaray, S. S. (2020). A comprehensive survey and analysis of generative models in machine learning. *Computer Science Review*, *38*(2020), 100285. https://doi.org/10.1016/j.cosrev.2020.100285

Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, *6*(1), 53–60. https://doi.org/10.21427/D7CF7R

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structural analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

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

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

Ilhan, M., & Cetin, B. (2014). Development of Classroom Assessment Environment Scale (CAES): Validity and reliability study. *Education and Science*, *39*(176), 31–50. https://doi.org/10.15390/EB.2014.3334

Jang, Y., Choi, S., Kim, H., & Development, 27. (2022). and validation of an instrument to measure undergraduate students' attitudes toward the ethics of artificial intelligence (AT-EAI) and analysis of its difference by gender and experience of AI education. *Education and Information Technologies*, *27*(8), 11635–11667. https://doi.org/10.1007/s10639-022-11086-5

Jeon, J., Lee, S., & Choi, S. (2023). A systematic review of research on speech-recognition chatbots for language learning: Implications for future directions in the era of large language models. *Interactive Learning Environments*, 1–19. https://doi.org/10.1080/10494820.2023.2204343

Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, *62*(1), 15–25. https://doi.org/10.1016/j.bushor.2018.08.004

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

Kelly, S., Kaye, S. A., & Oviedo-Trespalacios, O. (2023). What factors contribute to acceptance of artificial intelligence? A systematic review. *Telematics and Informatics*, *77*(2023), 101925. https://doi.org/10.1016/j.tele.2022.101925

Kline, R. B. (1994). *An easy guide to factor analysis*. Routledge.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). Guilford Press.

Küchemann, S., Steinert, S., Revenga, N., Schweinberger, M., Dinc, Y., Avila, K. E., & Kuhn, J. (2023). *Physics task development of prospective physics teachers using ChatGPT*. arXiv preprint arXiv:2304.10014.

Li, B., Kou, X., & Bonk, C. J. (2023). Embracing the disrupted language teaching and learning field: Analyzing YouTube content creation related to ChatGPT. *Languages*, *8*(3), 197. https://doi.org/10.3390/languages8030197

Li, Z., Shi, L., Wang, J., Cristea, A. I., & Zhou, Y. (2023). Sim-GAIL: A generative adversarial imitation learning approach of student modelling for intelligent tutoring systems. *Neural Computing and Applications*, *35*(34), 24369–24388. https://doi.org/10.1007/s00521-023-08989-w

Lieto, A., Bhatt, M., Oltramari, A., & Vernon, D. (2018). The role of cognitive architectures in general artificial intelligence. *Cognitive Systems Research*, *48*(2018), 1–3. https://doi.org/10.1016/j.cogsys.2017.08.003

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

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

Miles, J., & Shevlin, M. (2007). A time and a place for incremental fit indices. *Personality and Individual Differences*, *42*(5), 869–874. https://doi.org/10.1016/j.paid.2006.09.022

MIT Technology Review. (2022a). *Generative AI is changing everything. But what's left when the hype is gone?* https://www.technologyreview.com/2022/12/16/1065005/generative-ai-revolution-art/

MIT Technology Review. (2022b). *Get ready for the next generation of AI*. https://www.technologyreview.com/2022/10/03/1060561/get-ready-for-the-next-generation-of-ai/

Mitchell, M. (2019). *Artificial intelligence: A guide for thinking humans*. Penguin UK.

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

Pallant, J. (2005). *SPSS survival manual: A step by step guide to data analysis using SPSS for windows*. Allen & Unwin.

Park, Y. J. (2021). *The future of digital surveillance: Why digital monitoring will never lose its appeal in a world of algorithm-driven AI*. University of Michigan Press.

Park, Y. J., & Jones-Jang, S. M. (2022). Surveillance, security, and AI as technological acceptance. *AI & Society*, 1–12. https://doi.org/10.1007/s00146-021-01331-9

Prasad Agrawal, K. (2023). Towards adoption of generative AI in organizational settings. *Journal of Computer Information Systems*, 1–16. https://doi.org/10.1080/08874417.2023.2240744

Raffaghelli, J. E., Rodríguez, M. E., Guerrero-Roldán, A. E., & Baneres, D. (2022). Applying the UTAUT model to explain the students' acceptance of an early warning system in higher education. *Computers & Education*, *182*(2022), 104468. https://doi.org/10.1016/j.compedu.2022.104468

Roll, I., & Wylie, R. (2016). Evolution and revolution in artificial intelligence in education. *International Journal of Artificial Intelligence in Education*, *26*(2), 582–599. https://doi.org/10.1007/s40593-016-0110-3

Ruiz-Rojas, L. I., Acosta-Vargas, P., De-Moreta-Llovet, J., & Gonzalez-Rodriguez, M. (2023). Empowering education with generative artificial intelligence tools: Approach with an instructional design matrix. *Sustainability*, *15*(15), 11524. https://doi.org/10.3390/su151511524

Sallam, M., Salim, N., Barakat, M., Al-Mahzoum, K., Al-Tammemi, A. B., Malaeb, D., Hallit, R., & Hallit, S. (2023). Assessing attitudes and usage of ChatGPT in Jordan among health students: A validation study of the Technology Acceptance Model-Based Scale (TAME-ChatGPT). *JMIR Medical Education*, *9*(2023), e48254. https://doi.org/10.2196/preprints.48254

Sezer, B., & Yilmaz, R. (2019). Learning Management System Acceptance Scale (LMSAS): A validity and reliability study. *Australasian Journal of Educational Technology*, *35*(3), 15–30. https://doi.org/10.14742/ajet.3959

Shidiq, M. (2023, May). The use of artificial intelligence-based chat-gpt and its challenges for the world of education; from the viewpoint of the development of creative writing skills. *Proceeding of International Conference on Education, Society and Humanity* (Vol. 1, No. 1, pp. 353–357).

Sohn, K., & Kwon, O. (2020). Technology acceptance theories and factors influencing artificial Intelligence-based intelligent products. *Telematics and Informatics*, *47*(2020), 101324. https://doi.org/10.1016/j.tele.2019.101324

Stevens, J. (1996). *Applied multivariate statistics for the social sciences*. Erlbaum.

Strzelecki, A. (2023). To use or not to use ChatGPT in higher education? A study of students' acceptance and use of technology. *Interactive Learning Environments*, 1–14. https://doi.org/10.1080/10494820.2023.2209881

Sumer, N. (2000). Yapısal eşitlik modelleri [Structural equation models]. *Turkish Psychological Articles*, *3*(6), 49–74.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Pearson.

Tamilmani, K., Rana, N. P., Wamba, S. F., & Dwivedi, R. (2021). The extended Unified Theory of Acceptance and Use of Technology (UTAUT2): A systematic literature review and theory evaluation. *International Journal of Information Management*, *57*(2021), 102269. https://doi.org/10.1016/j.ijinfomgt.2020.102269

Tegmark, M. (2018). *Life 3.0: Being human in the age of artificial intelligence*. Vintage.

Teng, Z., Cai, Y., Gao, Y., Zhang, X., & Li, X. (2022). Factors affecting learners' adoption of an educational metaverse platform: An empirical study based on an extended UTAUT model. *Mobile Information Systems*, *Special Issue* (2022), 1–15. https://doi.org/10.1155/2022/5479215

Tezbaşaran, A. (1997). *Likert tipi ölçek hazırlama kılavuzu [Likert type scale preparation guide]*. Türk Psikologlar Derneği.

The New York Times. (2023). *A new area of A.I. booms, even amid the tech gloom*. https://www.nytimes.com/2023/01/07/technology/generative-ai-chatgpt-investments.html

Toh, S. Y., Ng, S. A., & Phoon, S. T. (2023). Accentuating technology acceptance among academicians: A conservation of resource perspective in the Malaysian context. *Education and Information Technologies*, *28*(3), 2529–2545. https://doi.org/10.1007/s10639-022-11288-x

Ustun, A. B., Karaoglan-Yilmaz, F. G., & Yilmaz, R. (2023). Educational UTAUT-based virtual reality acceptance scale: A validity and reliability study. *Virtual Reality*, *27*(2), 1063–1076. https://doi.org/10.1007/s10055-022-00717-4

Vărzaru, A. A. (2022). Assessing artificial intelligence technology acceptance in managerial accounting. *Electronics*, *11*(14), 2256. https://doi.org/10.3390/electronics11142256

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology. *MIS Quarterly*, *27*(3), 425–478. https://doi.org/10.2307/30036540

Venkatesh, V., Thong, J. Y., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly*, *36*(1), 157–178. https://doi.org/10.2307/41410412

Wang, H., & Wang, S. (2010). User acceptance of mobile internet based on the unified theory of acceptance and use of technology: Investigating the determinants and gender differences. *Social Behavior and Personality*, *38*(3), 415–426. https://doi.org/10.2224/sbp.2010.38.3.415

World Economic Forum. (2023). *What is generative AI? An AI explains*. https://www.weforum.org/agenda/2023/02/generative-ai-explain-algorithms-work/

Yilmaz, H., Maxutov, S., Baitekov, A., & Balta, N. (2023). Student attitudes towards Chat GPT: A technology acceptance model survey. *International Educational Review*, *1*(1), 57–83. https://doi.org/10.58693/ier.114

Yilmaz, R., & Karaoglan Yilmaz, F. G. (2023a). The effect of generative artificial intelligence (AI)-based tool use on students' computational thinking skills, programming self-efficacy and motivation. *Computers and Education: Artificial Intelligence*, *4*(2023), 100147. https://doi.org/10.1016/j.caeai.2023.100147

Yilmaz, R., & Karaoglan Yilmaz, F. G. (2023b). Augmented intelligence in programming learning: Examining student views on the use of ChatGPT for programming learning. *Computers in Human Behavior: Artificial Humans*, *1*(2), 100005. https://doi.org/10.1016/j.chbah.2023.100005

Yilmaz, R., Yurdugül, H., Karaoglan Yilmaz, F. G., Şahĭn, M., Sulak, S., Aydin, F., Tepgeç, M., Müftüoğlu, C. T., & Ömer, O. R. A. L. (2022). Smart MOOC integrated with intelligent tutoring: A system architecture and framework model proposal. *Computers and Education: Artificial Intelligence*, *3*(2022), 100092. https://doi.org/10.1016/j.caeai.2022.100092

About the authors

Fatma Gizem Karaoglan Yilmaz is an Associate Professor of Computer Technology and Information Systems at Bartin University. She is interested in distance education, flipped learning, interactive learning environments, human–computer interaction, virtual reality, augmented reality, and eye-tracking.

Ramazan Yilmaz is a Professor in the Department of Computer Technology & Information Systems at Bartin University in Turkey. He is interested in cyber psychology, smart learning environments, virtual reality and augmented reality, human–computer interaction, data mining, learning analytics, and eye tracking.

Mehmet Ceylan is an Instructor in the Department of Communication Coordinator at Bartin University in Turkey. He is the head of this department.

Appendix 1. Generative Artificial Intelligence Acceptance Scale

Factor	No.	Item (English)	Madde (Türkçe)	(1) Strongly disagree	(2) Disagree	(3) Neither agree or disagree	(4) Agree	(5) Strongly agree
Performance expectancy	1	I find generative AI applications useful in my daily life.	Üretken yapay zeka uygulamalarını günlük hayatımda faydalı bulurum.					
	2	The use of generative AI applications increase my chances of achieving the things that are important to me.	Üretken yapay zeka uygulamalarının kullanımı benim için önemli olan şeyleri başarma şansımı artırır.					
	3	Generative AI applications help me get things done faster.	Üretken yapay zeka uygulamaları işleri daha hızlı bitirmeme yardımcı olur.					
	4	Using generative AI applications increase my productivity.	Üretken yapay zeka uygulamalarının kullanımı benim üretkenliğimi arttırır.					
	5	The use of generative AI applications make my life easier.	Üretken yapay zeka uygulamalarının kullanımı hayatımı kolaylaştırır.					
	6	Generative AI applications are useful for my daily life.	Üretken yapay zeka uygulamaları günlük yaşamamım için kullanışlıdır.					
	7	The use of generative AI applications increase my chances of solving the problems I face.	Üretken yapay zeka uygulamalarının kullanımı karşıma çıkan problemleri çözme şansımı artırır.					
Effort expectancy	8	Learning how to use generative AI applications is easy for me.	Üretken yapay zeka uygulamalarını kullanmayı öğrenmek benim için kolaydır.					
	9	I think it is easy to leverage generative AI applications.	Üretken yapay zeka uygulamalarından yararlanmanın kolay olduğunu düşünüyorum.					
	10	Generative AI applications are easy to use.	Üretken yapay zeka uygulamalarını kullanmak kolaydır.					
	11	It is easy for me to become skilled in using generative AI applications.	Üretken yapay zeka uygulamalarını kullanma konusunda beceri sahibi olmak benim için kolaydır.					
	12	My interaction with generative AI applications is clear and understandable.	Üretken yapay zeka uygulamaları ile etkileşimim açık ve anlaşılırdır.					
Facilitating conditions	13	Generative AI applications are compatible with other technologies I use.	Üretken yapay zeka uygulamaları kullandığım diğer teknolojilerle uyumludur.					
	14	I can get help from others when I have difficulties in using generative AI applications.	Üretken yapay zeka uygulamalarının kullanımında zorluk yaşadığımda başkalarından yardım alabilirim.					
	15	If I experience any problems while using generative AI applications, I can access the necessary information for a solution.	Üretken yapay zeka uygulamalarını kullanırken herhangi bir sorun yaşarsam çözüme yönelik gerekli bilgilere ulaşabilirim.					
Social influence	16	People important to me think I should use generative AI applications.	Benim için önemli insanlar üretken yapay zeka uygulamalarını kullanmam gerektiğini düşünüyor.					
	17	The people I model my behavior on think I should use generative AI applications.	Davranışlarımda model aldığım kişiler üretken yapay zeka uygulamalarını kullanmam gerektiğini düşünüyor.					
	18	People whose opinions I value prefer me to use generative AI applications.	Düşüncelerine değer verdiğim kişiler üretken yapay zeka uygulamalarını kullanmamı tercih ediyorlar.					
	19	People who are important to me are using generative AI applications.	Benim için önemli insanlar üretken yapay zeka uygulamalarını kullanıyor.					
	20	People who are important to me encourage the use of generative AI applications.	Benim için önemli insanlar üretken yapay zeka uygulamalarını kullanmamı teşvik ediyor.					

If the acceptance of any generative AI application is to be examined, the name of the generative AI application can be written instead of "Generative AI applications" in the scale items.