INSTRUCTIONAL CHATBOT PRACTISE IN DISTANCE EDUCATION AND INVESTIGATION OF ITS ROLE IN THE LEARNING PROCESS

Dr. Ahmet Yusuf CEVHER
ORCID: 0000-0002-0233-3226
Department of Computer Engineering
Ardahan University
Ardahan, TURKIYE

Dr. Serkan YILDIRIM
ORCID: 0000-0002-8277-5963
Department of Computer Education and Educational Technologies
Ataturk University
Erzurum, TURKIYE

Received: 13/09/2024 **Accepted:** 25/02/2025

ABSTRACT

This study investigates the impact and role of an instructional chatbot, ARUChatbot, in a distance education setting. Using a sequential explanatory mixed-methods design, the research involved 130 students from Ardahan University's Basic Information Technologies course. Participants were selected through purposive sampling. Quantitative data were collected via achievement tests and log records, while qualitative data were obtained through focus group interviews. Findings indicate that ARUChatbot positively influences students' academic achievement. Students utilized the chatbot for activities such as information retrieval, completing assignments, class preparation, and content reinforcement. The findings indicate that ARUChatbot sees higher demand during busy periods like exam weeks and effectively supports students' assignment needs, with increased usage as deadlines approach. Qualitative data revealed that students appreciated the chatbot's instant response and concise information delivery. However, some students criticized its text-based nature and requested enhancements, including visual or auditory support and mobile access. The reluctance of some students to use the chatbot was linked to insufficient promotion and training. This study suggests that instructional chatbots like ARUChatbot can effectively support academic achievement and meet students' learning needs, particularly during high-demand periods. In the qualitative findings, the chatbot's features of providing immediate responses and delivering concise, relevant information were positively received by the students. However, the text-based interaction and certain limitations were criticized. Students have requested improvements such as visual or audio input-output support and mobile application accessibility.

Keywords: Artificial intelligence, instructional chatbot, academic success, mixed pattern, ARUChatbot.

INTRODUCTION

The transformation of technology in the field of education has become increasingly evident in recent times. The integration of technology as an alternative to traditional teaching methods enriches students' learning experiences and makes teaching processes more effective. This integration occurs through various platforms and tools. These platforms provide students with diverse learning experiences, allowing them to learn at their own pace, while also enabling teachers to monitor student progress more closely. Interactive applications and digital tools enhance students' engagement with the course content (Arslan et al., 2022). The use of technology in education is also crucial from the perspectives of accessibility and inclusivity. Technology can provide broader access to learning materials and increase the participation of disadvantaged individuals in education, thereby promoting educational equity (Ahmad, 2015). For example, distance education offers students the opportunity to overcome geographical, temporal, and other constraints, allowing them to receive education anytime and anywhere (Altwaijry et al., 2021; Gunawardena & McIsaac, 2013). This educational model is typically implemented through written materials, online platforms, video conferences, and other communication technologies. Nowadays, both academic institutions and corporate training sectors widely utilize distance education to provide flexibility and accessibility for students and employees. Furthermore, distance education offers opportunities to overcome geographical barriers, adapt to different learning styles, and better address students' individual needs (Pendy, 2023).

Distance education is one of the area's most significantly affected by technological transformations. The pandemic has accelerated technological changes in distance education (Balyer & Oz, 2018). However, there is an increasing need for technologies that provide pedagogical support to students in distance education. These technologies can include various tools designed to support students' learning processes, enhance their motivation, and assist them in overcoming learning difficulties (Edisherashvili et al., 2022; Zaikin et al., 2016). One of the key tools for providing pedagogical support in distance education is chatbots (Tsivitanidou & Ioannou, 2021). Chatbots can perform a range of functions, such as providing information on specific topics, assisting with assignments, answering learning-related questions, and even offering emotional support. By utilizing generative artificial intelligence, natural language processing, and machine learning techniques, chatbots can offer human-like interactions. Additionally, chatbots can present customized learning content to students and recommend materials that align with their interests and learning styles (Tsivitanidou & Ioannou, 2021; Wollny et al., 2021). Students can interact with chatbots to obtain information about learning materials, understand concepts, and resolve uncertainties in the learning process by asking questions. For instance, when a student needs help with a topic, chatbots can provide immediate responses, resolve the student's issues, or facilitate access to relevant resources. Furthermore, chatbots can offer feedback to students and assist in monitoring their learning progress (Ait Baha et al., 2023).

CHATBOTS

Chatbots have become a significant interface for interacting with humans, thanks to advancements in artificial intelligence and natural language processing (Reznik, 2009). These systems typically communicate via text or voice, answering users' questions, providing information, or engaging in casual conversation for entertainment purposes (Nuruzzaman & Hussain, 2018). The operational framework of chatbots generally relies on natural language processing (NLP) and machine learning techniques. Their architecture typically consists of components such as an automatic speech recognizer, a language understanding module, a dialogue manager, and a speech synthesis module (Nimavat & Champaneria, 2017). Initially, the user's input is processed, and then analyzed using modeling and algorithms to generate an appropriate response. The dialogue manager retrieves relevant data from the database and sends it to the speech generator to be converted into text. This process occurs within a continuously evolving loop (Shum et al., 2018).

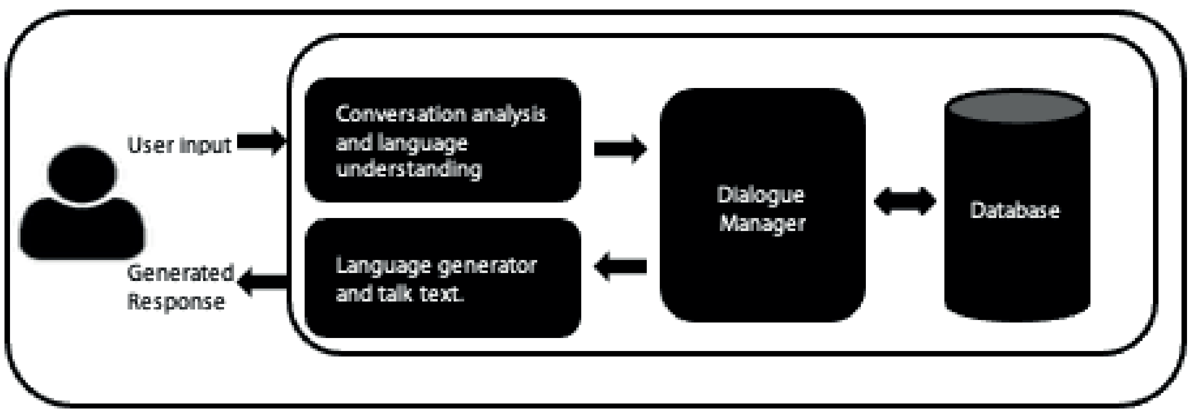

Figure 1. Working Principle of Chatbots

Chatbots have undergone significant evolution from their inception to the present day. A review of the historical development of chatbots in the literature reveals that their roots trace back to 1950, when a computer program was tested for its ability to interact with humans in a manner like human communication. Alan Turing's concept of the "Turing Test" proposed that computers could possess human-like intelligence and language understanding abilities during this period (Turing, 1950). However, due to the limited technology and information processing capabilities at that time, it was not possible to develop a fully functional chatbot. The communication capabilities of early chatbots, though limited, served as a guiding influence for the design of subsequent chatbots (Adamopoulou & Moussiades, 2020). From the 1960s onward, researchers began developing computer programs with basic language processing and conversational abilities. For instance, the ELIZA program, developed by Joseph Weizenbaum at MIT in 1966, was notable for its ability to conduct psychotherapy-like conversations (Weizenbaum, 1966). ELIZA could interact with users at a basic level and provide responses by manipulating the given inputs. Subsequently, in 1979, the PARRY program drew attention in an experiment designed to distinguish a real schizophrenic patient. PARRY, a chatbot that could mimic psychiatric disorders, was developed by Kenneth Colby from Stanford University (Heiser et al., 1979). The 1980s and 1990s saw significant advancements in artificial intelligence and natural language processing. During this period, computers became powerful enough to support more complex language processing algorithms, and researchers began developing more human-like chatbots using linguistic and psychological models (Balaceanu, 2020). In 1995, ALICE, the first online chatbot, was created. ALICE (Artificial Linguistic Internet Computer Entity), developed by Richard Wallace, was one of the first chatbots capable of conducting interactive conversations over the Internet, utilizing natural language processing algorithms to engage in meaningful dialogues with users (Wallace, 2009). With the rise of the Internet and digital technologies in the early 2000s, chatbots began reaching a broader audience. In 2001, ActiveBuddy developed the SmarterChild chatbot, which became the first chatbot to operate on a popular instant messaging platform, AOL Instant Messenger. SmarterChild could engage in interactive conversations with users, providing weather updates, news, and various other information (Molnar & Szuts, 2018). In 2011, Apple developed Siri, a chatbot that served iPhone users. Siri functioned as a virtual assistant capable of processing natural language questions and commands (Shum et al., 2018). In 2012, Google introduced Google Now (Google Assistant), leveraging Google's infrastructure to offer interactive services through voice commands on smartphones, smart speakers, and other Google products (Chandrasekar, 2023). In 2014, Amazon developed Alexa, a virtual assistant providing interactive services through natural language. Alexa, available on Amazon's smart speaker, the Amazon Echo, continues to perform tasks such as home automation, playing music, providing news, and more (Chen et al., 2018). One of the latest advancements in the field is ChatGPT, an artificial intelligence model developed by OpenAI. Released in 2022, ChatGPT is part of the GPT (Generative Pre-trained Transformer) family launched by OpenAI in 2018. ChatGPT is an AI model capable of interacting with users in natural language and covering a wide range of topics. This model has advanced natural language understanding and generation abilities through deep learning techniques (OpenAI, 2022). The historical development of chatbots from the past to the present is summarized in Figure 2.

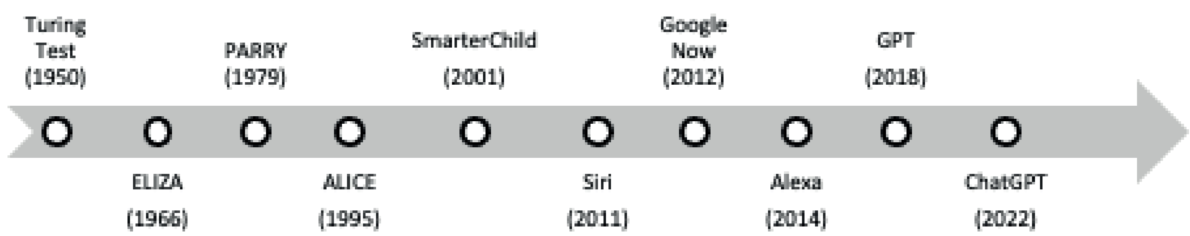

Figure 2. Historical Development of Chatbots

Today, advancements in artificial intelligence, machine learning, and deep learning are further enhancing the capabilities of chatbots, broadening their applications across various fields. It is anticipated that chatbots will become even more widespread and functional in the future. In particular, chatbots could play significant roles in the field of education, especially in remote learning processes and learning management systems (Verleger & Pembridge, 2018).

As artificial intelligence technologies continue to evolve, their potential applications in education are becoming increasingly apparent. While the general advancements in AI, machine learning, and deep learning have paved the way for the development of more sophisticated chatbots, their specific role in education demands closer examination. This growing interest has prompted numerous studies focusing on educational chatbots, exploring their potential to transform teaching and learning processes. (Adamopoulou & Moussiades, 2020; Hwang & Chang, 2021; Okonkwo & Ade-Ibijola, 2021; Winkler & Sollner, 2018; Wollny et al., 2021). Early studies (Winkler & Sollner, 2018; Hobert & Wolff, 2019; Wollny et al., 2021) explored their application in healthcare and language learning, emphasizing their early developmental stage and superficial examination in the literature. Additionally, many researchers highlighted the predominance of mobile interfaces in chatbot design (Smutny & Schreiberova, 2020; Jeon et al., 2023).

More recent works focus on the potential of chatbots for specific educational domains. Huang et al. (2022) and Lo (2023) addressed chatbots' impact on language learning, highlighting benefits such as increased social presence alongside concerns about accuracy and plagiarism. Similarly, Imran and Almusharraf (2023) investigated plagiarism risks with ChatGPT, advocating for a re-evaluation of assessment methods in academic contexts.

The advent of ChatGPT has further fueled interest in the field, with numerous studies exploring its applications. Research has examined its role in language learning (Kohnke, 2023; Mohamed, 2023; Muniandy & Selvanathan, 2024; Yan, 2023; Zadorozhnyy & Lai, 2023) and other educational areas, such as science, medicine, and psychology (Clark, 2023; Keiper et al., 2023; Liu et al., 2024; Polverini & Gregorcic, 2024). Its ethical implications, security concerns, and broader educational impact remain ongoing topics of discussion (Crawford et al., 2023; Markowitz, 2024; Nikolic et al., 2023).

Recent contributions have expanded the discussion to include the integration of chatbots into pedagogical frameworks. Ahn et al. (2024) explored ChatGPT's dual role as a disruptor and a teaching assistant in English Language Teaching (ELT). Bai et al. (2024) conducted a comparative study on chatbot-supported instructional scaffolding versus human teaching, revealing promising outcomes for learning performance. Yusuf et al. (2025) proposed a conceptual framework for AI conversational agents in higher education, highlighting their potential for personalized learning.

Another focus area is student engagement and self-regulated learning. Lai (2024) investigated how generative AI chatbots facilitate self-regulation in students. Similarly, Yin et al. (2024) analyzed the interplay between chatbot-induced emotions and students' learning motivation. Ng et al. (2024) provided insights into using ChatGPT for promoting self-regulated science learning, showcasing its transformative potential in education.

Despite these advancements, there remains a lack of in-depth studies on the interactions between students and chatbots, including the messages exchanged and the contextual factors influencing chatbot usage. Furthermore, there is limited research on how students' tendencies to use chatbots evolve based on their educational environment and learning needs. Addressing these gaps is crucial for better understanding the integration of chatbots into learning processes and maximizing their contributions to education.

PURPOSE OF THE STUDY

This study aims to reveal the potential impact and role of an instructional chatbot designed for use in remote education. The research questions guiding this study are as follows:

1. What is the effect of instructional chatbot usage on academic achievement?
   1.1. Does academic achievement vary according to the use of the instructional chatbot?
   1.2. Is there a statistically significant relationship between the number of messages written by students and their academic achievement?

2. What types of learning activities do students write messages about? What is the distribution of messages written by students concerning learning activities?
3. What are the students' tendencies towards using instructional chatbots?
    3.1. How is the temporal distribution of instructional chatbot usage?
    3.2. Does the use of instructional chatbots vary according to assignment deadline?
4. What are students' experiences with using chatbots?

This study examines the effective use of instructional chatbots in education and their potential to enhance student success. By investigating the impact of instructional chatbots on learning processes, it highlights the importance of technology-based innovative applications. Additionally, by exploring the effects of instructional chatbots on various variables, the study aims to provide a foundation for future research. According to the literature, the effect of instructional chatbots on academic achievement has not been thoroughly examined. This study addresses this gap by evaluating features related to chatbot usage, such as message count and temporal aspects. Furthermore, it examines instructional chatbot use in relation to temporal variables such as assignment submission and exam days. By focusing on the learning activities conducted with the instructional chatbot and their cognitive levels, the study aims to demonstrate its effects on learning. Finally, the study provides recommendations on the effects of instructional chatbots on academic success and learning processes by considering users' views on design and implementation.

METHOD

Research Model

In this study, which aims to uncover the potential impact and role of an instructional chatbot (ARUChatbot) designed for use in remote education, a sequential explanatory design from mixed research methods has been employed. In this design, the research is primarily conducted quantitatively and supported by qualitative data. The combined or sequential use of qualitative and quantitative research methods is considered more effective than using them separately (Creswell, 2013). This design helps researchers answer research questions more comprehensively and can overcome the limitations of quantitative or qualitative research designs when used alone (Creswell, 2013). An overview of the research methods used in alignment with the research questions of the study is presented in Figure 3.

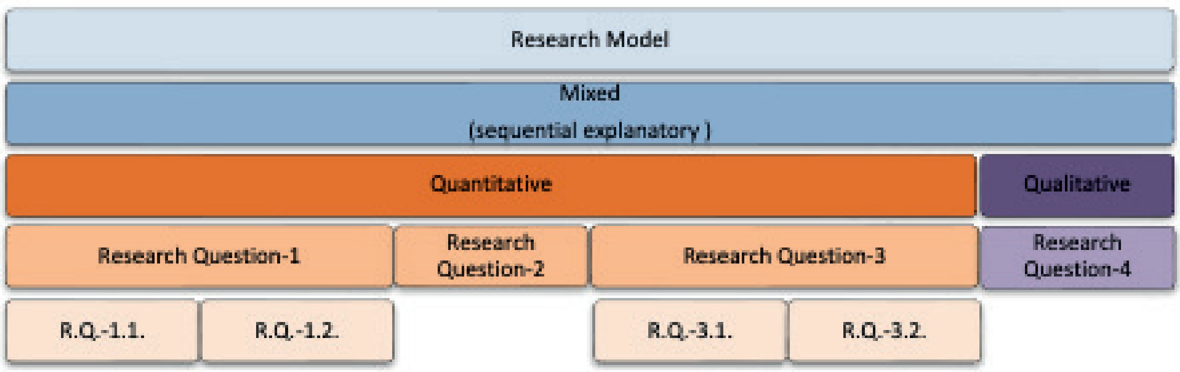

Figure 3. Research Methods Overview

In the quantitative part of the study, a pre-test/post-test control group quasi-experimental design has been used. In this research design, two groups (experimental and control) are formed to evaluate the effects of the intervention. However, only one group (the experimental group) receives the intervention, and measurements are taken before (pre-test) and after (post-test) the intervention in both groups. This allows for a clearer observation of the effects of the intervention (Ozmen, 2014). The quasi-experimental design used in the quantitative part of the study is detailed in Table 1.

Table 1. Pre-Test/Post-Test Control Group Design

Group	Pre-Test	Intervention	Post-Test
Experimental group	O1,1	X	O1,2
Control group	O2,1		O2,2

In the quantitative part of the study, students in the experimental group were provided with a technological tool (ARUChatbot) for use outside of class hours. Additionally, these students were given weekly assignments and informed that their assignments would be graded. They were expected to use the chatbot to assist with completing these assignments. In contrast, students in the control group were also given weekly assignments, but no technological tools were provided to assist them.

In the qualitative part of the study, a case study design has been employed. A case study is a qualitative research method aimed at conducting an in-depth examination of a particular individual, group, event, organization, or community. It explores how factors related to the case affect the situation or how the situation impacts these factors. Case studies not only describe the case being studied but also seek to explain cause-and-effect relationships and processes (Yildirim & Simsek, 1999).

Participants

The participants in the study consist of 130 students enrolled in the Basic Information Technologies course at Ardahan University, which is delivered through remote education. The participants were selected using a purposive sampling method, specifically homogeneous sampling. This sampling method involves choosing participants who possess certain characteristics or experiences relevant to the research's goals or objectives (Buyukozturk et al., 2008). Due to the constraints in distinguishing students on an individual basis under the current conditions, participants were grouped based on their academic programs. When establishing the experimental and control groups, considerations included the average scores, number of students, and gender distribution within the programs to form these groups. Information about the participants is presented in Table 2.

Table 2. Table of Participants

Group	Department	Gender		Pre-Test Score Average
		Male	Female	
Experimental group	Construction	5	5	34.43
	Occupational Health and Safety	15	12	
	Health Information Systems Technician	13	11	
	Total:	33	28	
Control group	Electrical	20	1	35.81
	Opticianry	4	21	
	Hair Care and Beauty Services	5	18	
	Total:	29	40	

As seen in Table 2, the experimental group comprises 33 male and 28 female students, totaling 61 students, while the control group consists of 29 male and 40 female students, totaling 69 students. After the groups were formed, an Independent Samples T-Test revealed no significant differences between the groups. Table 3 provides the number of samples involved in the data collection processes according to the research questions.

Table 3. Number of Participants According to Research Questions

Research Design	Research Question		Number of Participants
Quantitative	R.Q.-1	R.Q.-1.1.	130
		R.Q.-1.2.	61
	R.Q.-2		61
	R.Q.-3	R.Q.-3.1.	49
		R.Q.-3.2.	49
Qualitative	R.Q.-4		15

Data Collection Tools

In the study, data were collected using pre-tests and post-tests for assessing achievement, log records documenting students' interactions with ARUChatbot, and a semi-structured interview form. Table 4 provides details on the data collection tools used to gather data according to the research questions.

Table 4. Data Collection Tools According to Research Questions

Research Design	Research Question		Data Collection Tools
Quantitative	R.Q.-1	R.Q.-1.1.	Achievement Test
		R.Q.-1.2.	Achievement Test and Log Records
	R.Q.-2		Log Records
	R.Q.-3	R.Q.-3.1.	Log Records
		R.Q.-3.2.	Log Records
Qualitative	R.Q.-4		Semi-Structured Interview Form

Achievement Test for Pre-Test

In the study's application to participants, the course materials from Anadolu University's free-access Basic Information Technologies course (Aydin et al., 2015, Sections 1-5) were utilized to create a pool of 40 pre-test items. To develop this pool, exam questions from the same course module were adapted. Before creating the item pool, learning objectives were defined, and a table of specification with main themes and sub-themes was prepared. Questions were then generated to ensure a balanced distribution across each main theme and sub-theme, as outlined in the table of specification. The pre-test form was reviewed by a Turkish Language and Literature faculty member with a doctoral degree to assess language and clarity. Based on the expert's feedback, revisions were made to certain items. To ensure content validity, feedback on the items was obtained from eight experts with various degrees (5 Ph.D., 1 Master's, and 2 Bachelor's) in Computer Science, using a three-option form: "suitable," "needs modification," or "not suitable." Content validity was assessed using the Lawshe content validity ratio (CVR), with a ratio of 0.750 determined for the eight experts (Lawshe, 1975). Following expert feedback, seven items with a CVR below 0.750 were removed, and four items were revised. The content validity of the remaining 37 items was recalculated, resulting in a CVR of 0.891, which is greater than 0.750, confirming the content validity of the 37-item test. Before applying the pre-test, a pilot test was conducted with five associate degree students in a face-to-face setting with an examiner to confirm the test's clarity. The final test was administered to 135 students (70 females and 65 males) across six different programs at Ardahan University Technical Sciences Vocational School. After the application, item difficulty and discrimination analyses were conducted using SPSS with the Simple Method. The analyses revealed that items 12, 20, 25, 32, and 36 had discrimination indexes of 0.19 or below, while item 30 was identified as negatively discriminative. Consequently, these six items were removed from the test. For the remaining 31 items, difficulty analysis showed that 3 items had difficulty indexes of 0.29 or below (difficult), 16 items fell in the range of 0.30 – 0.49 (medium difficulty), 14 items ranged from 0.50 – 0.69 (easy), and 1 item had a difficulty index of 0.70 – 1 (very easy). As a result, items 2, 18, and 21 (difficult) and items 24 and 30 (very easy) were removed from the test. The final 27-item test had an overall difficulty level of 0.48 (medium) and an overall discrimination index of 0.33 (rather good).

Achievement Test for Post-Test

In preparing the final test for the study, the same steps used for the pre-test were repeated. Initially, a specification table was created, and based on this table, a pool of 40 items was developed. The draft of the final test was reviewed by a language expert for clarity and comprehensibility, and modifications were made based on the expert's suggestions. Subsequently, the content validity of the final test was evaluated by 8 field experts using a form with three options: appropriate, needs revision, and not appropriate. The content validity was assessed using the Lawshe method, resulting in a content validity ratio (CVR) of 0.750 for the 8 experts. Following expert feedback, 8 items were found to be below the CVR threshold; 6 items were removed, and 2 items were revised. The content validity of the remaining 34 items was recalculated, showing a CVR of 0.911, which was above the threshold of 0.750. Thus, the 34-item test was deemed valid for use. Before administering the final test, a pilot test was conducted with 5 associate degree students to confirm its clarity. The test was then administered to 130 students (67 females and 63 males) from 6 different programs at Ardahan University Technical Sciences Vocational School. However, 3 students who took the pre-test did not participate in the final test. After administration, item difficulty and discrimination analyses were conducted using SPSS. The analysis revealed that items 16, 17, 21, 23, 24, 26, 27, 28, and 29 had discrimination indexes of 0.19 or below. Consequently, these 9 items were removed from the test. For the remaining 25 items, difficulty analysis showed that 2 items had difficulty indexes of 0.29 or below (difficult), 13 items were in the range of 0.30 – 0.49 (medium difficulty), 7 items were in the range of 0.50 – 0.69 (easy), and 3 items were in the range of 0.70 – 1 (very easy). Based on this analysis, the difficult item 20 and the very easy items 1, 2, and 3 were removed. The final 21-item test had an overall difficulty level of 0.45 (medium) and an overall discrimination index of 0.34 (rather good).

Interview Form

A semi-structured interview form was developed to gather users' opinions on the ARUChatbot following the final test. Initially, themes were identified, and questions were formulated to elicit data related to these themes. The interview form was reviewed by four subject matter experts, and based on their feedback, the themes were finalized. The interview form includes four themes for users of the chatbot: (i) Role in the learning process, (ii) Impact on success, (iii) Usage tendencies, and (iv) Usability. For non-users of the chatbot, there is one theme: (i) Reasons for non-use. Subsequently, a pilot study was conducted using the interview form with four volunteer students from the experimental group. Following the pilot study, the sub-questions related to the themes were revised to prevent data loss.

Data Analysis

Since the study was conducted using a mixed-methods design, it includes both quantitative and qualitative data. Therefore, various quantitative and qualitative analysis methods were employed in the analysis of the collected data. Table 5 presents the analysis methods used according to the research questions and data collection instruments.

Table 5. Data Analysis Methods According to Research Questions

Research Question		Data Collection Tools	Data Analysis Methods
R.Q.-1	R.Q.-1.1.	Achievement Test	Descriptive Statistics Normality Test Levene Homogeneity Test Independent Sample T-Test
	R.Q.-1.2.	Achievement Test and Log Records	Normality Test Spearman's Rank-Order Correlation
R.Q.-2		Log Records	Descriptive Statistics

R.Q.-3	R.Q.-3.1.	Log Records	Descriptive Statistics
	R.Q.-3.2.	Log Records	Post Hoc Tamhane's T2 Test Descriptive Statistics
R.Q.-4		Semi-structured interview form (Focus group)	Content Analysis

As shown in Table 5, quantitative analyses were used for the first, second, and third research questions, while qualitative analysis was preferred for the fourth research question. Descriptive statistics were employed to characterize the data descriptively, and an Independent Samples T-Test was used to compare the success levels between two different groups. For the assumptions of the Independent Samples T-Test, normality tests were conducted using the Shapiro-Wilk Test and measures of Skewness-Kurtosis. Levene's Test was used to assess the homogeneity of variances, another assumption requirement. Since the data for the second sub-question of the first research question did not follow a normal distribution, the non-parametric Spearman's Rank-Order Correlation Test was used. To explore students' tendencies to use the instructional chatbot in the third research question, descriptive statistics were employed to present the situation, and the Kruskal-Wallis H Test was used to examine variations between variables. To identify the source of the differences, the non-parametric Post Hoc Tamhane's T2 Test was applied. For the qualitative data analysis, content analysis was utilized. Content analysis categorizes qualitative data into themes determined by similarity, which simplifies the evaluation process and renders the data meaningful (Yildirim & Simsek, 1999).

Validity and Reliability

The validity and reliability of a study provide information about the accuracy and dependability of the study's fundamental arguments and results. Validity of the study indicates that the examined subject has been measured and analyzed accurately. It is related to the proper use of the research methods and data (Karakoc & Donmez, 2014; Yildirim & Simsek, 1999). Reliability, on the other hand, refers to the repeatability and consistency of the research process and results. It reflects the stability of the research process and the measurement tools used (Sencan, 2005). The research questions, which outline the purpose of the study, have been clearly defined. Methods, data collection tools, and data analysis techniques have been selected based on a thorough review of the literature to align with the research questions. Throughout the research, the opinion of an expert researcher has been sought at every step, ensuring peer review at each stage of the study. Each step of the research has been presented clearly and explicitly, providing a detailed description for researchers working in the field. When developing data collection tools, attention was given to ensuring internal consistency by consulting a language expert and eight field experts. A comprehensive data source was aimed to be created using various tools such as semi-structured interview forms, audio recordings, and log files in the qualitative data collection process. Pilot applications of the developed data collection tools were conducted. Additionally, for the quantitative data collection tools, validity, item discriminability, and item difficulty analyses were performed. In the qualitative section of the study, to ensure confirmability and credibility, categorization was performed in conjunction with an expert, and it was observed that the interpretations were consistent. In cases where interpretations did not overlap, the previous categorization was discarded, and re-categorization was performed. Although a consistency coefficient was not calculated in the study, coding was conducted through discussion, reaching a consensus (Watts & Finkenstaedt-Quinn, 2021). For qualitative data, the number of participants was increased until data saturation was reached, and data collection was concluded when redundancy began to appear. Additionally, to adhere to ethical principles, coding was used instead of students' names in the findings related to qualitative data. The research process was conducted in accordance with ethical principles, and necessary permissions were obtained from relevant institutions and organizations.

Design and Implementation of the Instructional Chatbot

Based on a review of the relevant literature, the features of similar chatbots used in educational studies were analyzed, and the design principles and operational features of the instructional chatbot to be developed were determined. Accordingly, a prototype named ARUChatbot was developed. ARUChatbot was designed as a closed-system chatbot with predefined boundaries and scope, supporting learning processes outside regular class hours by providing informative assistance.

Course-specific documents were uploaded to ARUChatbot with topic restrictions defined by the instructor. This allowed ARUChatbot to answer students' course-related questions using these documents while directing non-course-related questions to predefined responses. After the pilot implementation of ARUChatbot, several design changes were made. The chatbot's purpose shifted from an information-based structure to a response-generation-based model. Additionally, the chatbot model was updated from an artificial intelligence system to a generative artificial intelligence system. The final design features of ARUChatbot are summarized in Table 6.

Table 6. Features of ARUChatbot

Chatbot Information	Design Features
Chatbot Name	ARUChatbot
Social Presence	Robot
Purpose	Response-generation-based model
Response Generation	Retrieval-based model
Chatbot Model	Generative AI system
AI System Infrastructure	Natural Language Understanding (NLU)
Channel	Web (https://ARUChatbot.com/)
Communication Mode	Text-Based

Users created accounts on ARUChatbot by entering their email, password, first and last name, gender, and department information. After creating an account, users logged in by providing their email and password to access the chatbot. Outside class hours, users could freely interact with ARUChatbot through a chat interface. Messages from the chatbot and the users were displayed in different colors, with timestamps provided below each message. Additionally, like and dislike buttons were included to allow users to evaluate the chatbot's responses. ARUChatbot was introduced to students during live online classes conducted via distance education. Detailed information about the chatbot's usage was provided by the instructor. The chatbot's access link was shared during live sessions and on the course page within the university's information management system.

The experimental process was carried out in the Basic Information Technologies course, held synchronously every Monday at 6:00 PM via distance education. Weekly assignments were given to students to complete outside of class, with submission deadlines set for Sundays at 11:00 PM. In live sessions, students were regularly informed about the platform for submitting assignments, file formats, and deadlines. Additionally, the date of the final assessment exam was announced two weeks in advance. ARUChatbot was made available to students for a total of 39 days over the six-week course period. During this time, students confirmed that they did not use any other chatbots.

FINDINGS

This study is derived from the doctoral thesis titled "Instructional Chatbot Practice in Distance Education and Investigation of Its Role in the Learning Process," conducted at the Department of Computer and Instructional Technologies Education, Institute of Educational Sciences, Ataturk University.

RQ-1: What Is the Effect of Instructional Chatbot Usage on Academic Achievement?

RQ-1.1: Does Academic Achievement Vary According to The Use of The Instructional Chatbot?

To determine whether students' academic success varied based on the use of instructional chatbots, a comparison was made between the pre-test and post-test scores of students who used ARUChatbot and those who did not.

Findings regarding pre-test scores of experimental and control groups

To determine whether the experimental and control groups were equivalent before the implementation of ARUChatbot, a pre-test was administered to both groups using a success test. An Independent Samples T-Test was then used to analyze the results. For the Independent Samples T-Test to provide accurate results, two critical assumptions must be met: the data must exhibit normal distribution, and the group variances must be homogeneous (Kim & Park, 2019). The results of the normality test assessing the distribution of the data are presented in Table 7.

Table 7. Normality Values Regarding Pre-Test Scores of Groups

Groups	Shapiro-Wilk		Skewness	Kurtosis
	sd	p		
Experimental group	61	.188	.181	-.644
Control group	69	.150	-.247	-.325

According to the information presented in Table 7, the Shapiro-Wilk Test results for both groups are greater than .05 ($p \geq .05$), indicating that the data in both groups exhibit a normal distribution. Another important indicator of normality is that the skewness and kurtosis values fall within the range of +1.5 and -1.5 (Tabachnick & Fidell, 2013). After confirming that one of the assumptions for the Independent Samples T-Test is met, the Levene's Test results were examined to check whether the group variances are homogeneous. These results are presented in Table 8.

Table 8. Levene Test Results Regarding Pre-Test Scores of Groups

Levene Statistics	sd	p
.019	128	.889

According to the information presented in Table 8, the result of the Levene's Test is greater than .05 ($p \geq .05$), indicating that there is no significant difference between the variances of the experimental and control groups, meaning that the variances are homogeneous (Gastwirth et al., 2009). After confirming that the necessary assumptions are met, an Independent Samples T-Test was applied to compare the pre-test scores of the groups. The results of this test are presented in Table 9.

Table 9. Independent Sample T-Test Results on Pre-Test Scores of Groups

Groups	n	$\bar{X}$	ss	t test		
				t	sd	p
Experimental group	61	34.43	9.21	-.850	128	.397
Control group	69	35.81	9.32			

According to the data presented in Table 9, there is no statistically significant difference between the pre-test scores of the experimental and control groups (t[128] = -.850; p > .05). This result indicates that prior to the experimental intervention, there was no difference in the academic achievement levels of students in both the experimental and control groups, suggesting that the groups were comparable in terms of academic performance.

Findings regarding the post-test scores of the experimental and control groups

To compare the post-test scores of the experimental and control groups after the experimental intervention, the post-test scores were analyzed using the Independent Samples T-Test. Before conducting the Independent Samples T-Test, the assumptions of normal distribution of data and homogeneity of group variances were examined. The results of the normality test conducted to assess the normal distribution of the data are presented in Table 10.

Table 10. Normality Values Regarding the Post-Test Scores of the Groups

Groups	Shapiro-Wilk		Skewness	Kurtosis
	sd	p		
Experimental group	61	.073	.207	-.495
Control group	69	.238	.268	.064

According to the information provided in Table 10, the Shapiro-Wilk test results for both groups indicate p-values greater than .05 (p ≥ .05), and the skewness and kurtosis values fall within the range of +1.5 to -1.5. These results suggest that the data in both groups follow a normal distribution. The results of the Independent Samples T-Test, conducted to compare the post-test scores of the experimental and control groups, are presented in Table 11.

Table 11. Independent Sample T-Test Results on Post-Test Scores of Groups

Groups	n	$\bar{X}$	ss	t test		
				t	sd	p
Experimental group	61	51,78	19,46	3,153	99,983	,002
Control group	69	42,59	12,49			

According to the information presented in Table 11, there is a statistically significant difference between the post-test scores of the experimental and control group students (t [99.983] = -3.153; p < .05). This result indicates that after the experimental intervention, there is a meaningful difference in the academic achievement levels of students in the experimental and control groups, with students using ARUChatbot demonstrating higher academic success compared to those in the control group.

RQ-1.2: Is There a Statistically Significant Relationship Between the Number of Messages Written by Students and Their Academic Achievement?

Findings regarding the post-test scores of the experimental group students according to the number of messages written

A correlation analysis was conducted to determine the relationship between the number of messages written by the students in the experimental group and their academic scores. To decide which correlation coefficient should be used, the normality of the data was first examined. The results of the normality test for the distribution of the data are presented in Table 12.

Table 12. Normality Values Regarding the Number of Messages of Experimental Group Students

	Shapiro-Wilk		Skewness	Kurtosis
	sd	p		
Number of Messages	61	.000	2.804	9.168

According to the information provided in Table 12, the result of the Shapiro-Wilk test for the number of messages shows that the p-value is less than .05 ($p \leq .05$), and the skewness and kurtosis values are not within the acceptable range of +1.5 to -1.5. This indicates that the data do not follow a normal distribution. Therefore, the Spearman Rank-Order Correlation, which is appropriate for non-parametric data, was used to analyze the relationship between the number of messages written by the students in the experimental group and their academic scores (Field, 2013). The results of the correlation analysis are presented in Table 13.

Table 13. The Relationship Between Number of Messages and Academic Success

		Post-Test Score
Number of Messages	Spearman's r	.51*
	p	.00
	n	61

* The correlation is significant at p<.01 level.

According to Table 13, a strong positive ($r = .51$) and statistically significant ($p < .05$) correlation was found between the number of messages written and academic achievement. This indicates that the number of messages written by students in the experimental group is strongly and significantly associated with an increase in their academic performance. In other words, as the number of messages increases, academic success also increases.

RQ-2: What Types of Learning Activities Do Students Write Messages About? What Is the Distribution of Messages Written by Students Concerning Learning Activities?

Each message written by students to the instructional chatbot regarding course topics was considered a learning activity. To determine which learning activity each message corresponded to, log records were analyzed using content analysis, revealing four distinct categories. The number of messages for each of these categories is presented in Figure 4.

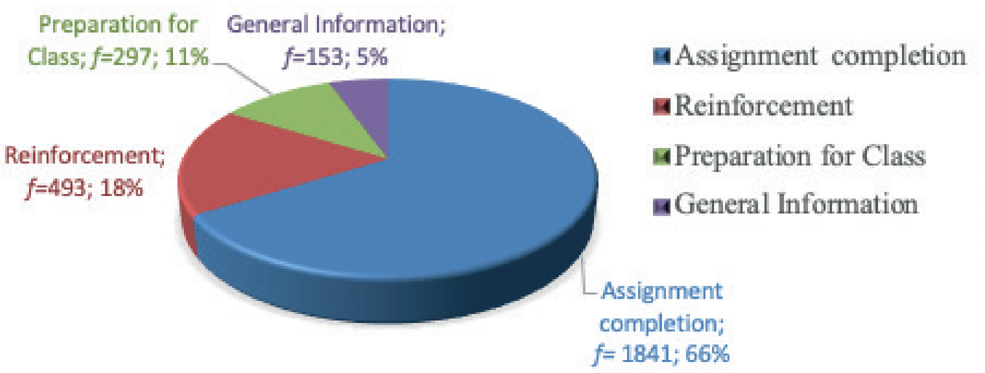

Figure 4. Distribution of Learning Activities by Message Count

Based on the information presented in Figure 4, it is evident that students most frequently used ARUChatbot for assignment completion. Among the messages written by students under the learning activity category, 66% were for completing assignments, 18% were for reinforcing information learned in class, 11% were for obtaining preliminary information about the next class topic to prepare for the lesson, and finally, 5% were for general information related to course topics. (i) Assignment Completion: This category includes messages related to assignments given weekly in the Basic Information Technologies course. Kvale and Brinkmann (2009) found that chatbots facilitate rapid access to information for students while completing assignments, thus helping in the effective completion of the tasks. (ii) Reinforcement: This category encompasses messages aimed at revisiting and reinforcing previously learned material to ensure retention. Chatbots can assist students in reinforcing and reviewing learned information. Fadhil (2018) noted that chatbots are effective in helping students reinforce knowledge and enhance retention. (iii) Preparation for Class: This category consists of messages written to obtain preliminary information about the upcoming week's topic to prepare for the lesson. Chatbots can be used to provide students with necessary pre-class information. Winkler and Sollner (2018) demonstrated that chatbots can be beneficial in helping students come better prepared for classes. (iv) General Information: This category includes messages written to seek general information about course topics outside of the assigned tasks. Chatbots can guide students in better understanding course materials and provide instant feedback, facilitating information retrieval on course-related topics.

RQ-3: What Are the Students' Tendencies towards Using Instructional Chatbots?

RQ-3.1: How Is the Temporal Distribution of Instructional Chatbot Usage?

Distribution of instructional chatbot usage by day and week

The usage distribution of ARUChatbot across different days, based on the number of messages, is presented in Figure 5. This analysis shows how frequently ARUChatbot was used on each day.

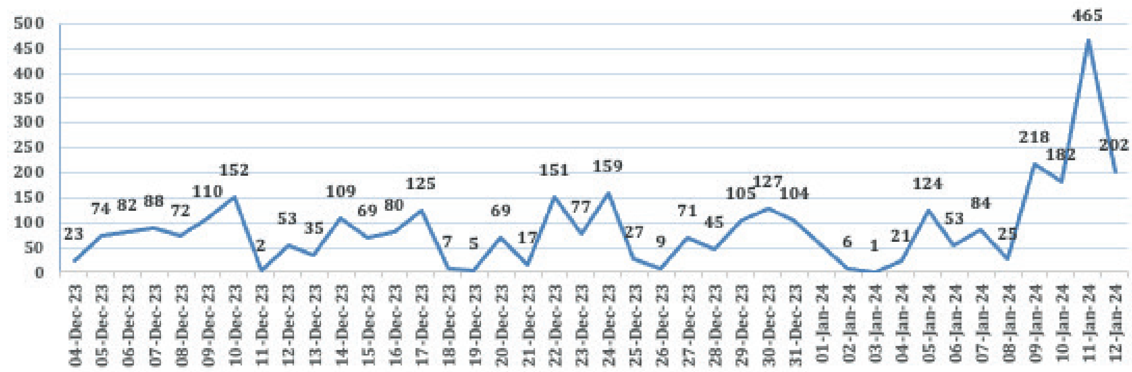

Figure 5. Distribution of ARUChatbot Usage by Day

The usage of ARUChatbot across weeks, based on the number of messages, is analyzed, and presented in Figure 6. This analysis illustrates how the usage of the instructional chatbot varied week by week.

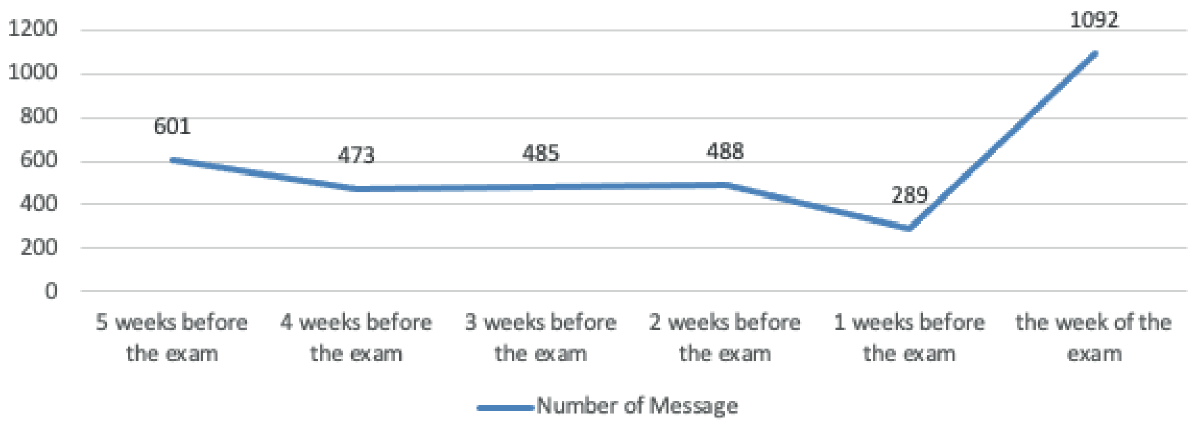

Figure 6. Distribution of ARUChatbot Usage by Week

Based on the information provided in Figure 6, the usage of ARUChatbot shows a declining trend after the first week but reaches its highest value during the exam week.

RQ-3.2: Does the Use of Instructional Chatbots Vary According to Assignment Deadline?

Distribution of instructional chatbot usage according to the assignment deadline

According to the information provided in Table 11, since the distribution of students' message counts is not normal, the Kruskal-Wallis H test, a non-parametric test, was used to examine whether the usage of the instructional chatbot varies according to the assignment deadline. The results of this test are presented in Table 14.

Table 14. Kruskal Wallis-H Test Results Showing Whether the Number of Messages Differs According to the Assignment Deadline

Score	Number of days left until the assignment deadline	N	Rank Average	Kruskal Wallis-H		
				X2	sd	p
Number of Messages	6 days before	5	7.40	14.56	6	.024
	5 days before	6	14.25			
	4 days before	6	17.75			
	3 days before	6	20.00			
	2 days before	6	27.08			
	1 days before	5	23.70			
	day of submission	5	30.00			

According to the information provided in Table 14, the Kruskal-Wallis H test indicated that there is a statistically significant difference between the ranks of the assignment deadline groups ($X^2 = 14.56$; df = 6; $p < .05$). To determine which specific groups contributed to this significant difference, a non-parametric Post Hoc Tamhane's T2 test was conducted (Hochberg & Tamhane, 1987). The results of this test are presented in Table 15.

Table 15. Post Hoc Tamhane's T2 Test Results Showing the Source of Differentiation

Number of days left until the assignment deadline (i)	Number of days left until the assignment deadline (j)	Rank Average (i-j)	p
6 days before	5 days before	-44.03333	.998
	4 days before	-56.53333	.795
	3 days before	-107.36667	.986
	2 days before	-103.70000	.069
	1 days before	-72.60000	.064
	day of submission	-108.00000*	.017

According to Table 15, the Post Hoc Tamhane's T2 Test revealed a statistically significant difference between the number of messages written 6 days before the assignment deadline and on the day of submission, with a higher number of messages on the submission day ($p < .05$). This indicates that the number of messages written on the submission day was significantly greater than the number of messages written 6 days before the deadline.

RQ-4: What Are Students' Experiences with Using Chatbots?

In the analysis of log records for participants selected from the experimental group, it was found that 9 students used ARUChatbot, 3 students completed their registration but wrote only one message or none, and 3 students did not register with ARUChatbot at all. The data collected from interviews were categorized into five main themes: learning process, usage frequency trends, usability, new ideas, and reasons for non-use. In these data, "*n*" represents the number of students, while "*f*" indicates the frequency of multiple occurrences.

The Impact of ARUChatbot on the Learning Process and Usage Frequency Trends

(i) Impact on the learning process

The impact of ARUChatbot on the learning process has been analyzed under two main headings:

Students primarily used ARUChatbot for activities such as completing assignments (9 students), acquiring information (6 students), preparing for classes (3 students), and reinforcement (3 students). Examples of student feedback:

E2: *…I used ARUChatbot for preparing assignments, presentations, and projects.*

E8: *…I used it to complete my assignments and get accurate answers to questions I asked in other subjects.*

E4: *…I asked about topics I was curious about, such as quantum physics.*

E3: *…I used it to get summary information about upcoming topics for preparation before classes.*

All students stated that ARUChatbot positively impacted their exam performance and contributed to their learning process. Examples of student feedback:

S3: *…I had a course I previously failed. I studied for that course's exam with ARUChatbot and passed.*

S4: *…It helped me discover new things. The more questions I asked, the more permanent my learning became.*

S6: *…Thanks to ARUChatbot, I achieved high grades in joint courses conducted via distance education.*

(ii) Usage frequency trends

The usage frequency of ARUChatbot varied depending on specific time frames, such as assignment submission deadlines and exam weeks. Most students stated that they used ARUChatbot more frequently on assignment submission days. Some preferred to use the chatbot before the submission date. Examples of student feedback:

S3: *…I used it more intensively on Sundays because I usually left my assignments to the last day.*

S9: *…I started using it two or three days before the assignment deadline, just in case there were any system issues...*

During exam weeks, students actively used ARUChatbot to support their exam preparation. Examples of student feedback:

S6: *…I used it more frequently before exam days to prepare for exams.*

S8: *…During exam weeks, I used ARUChatbot to find answers to questions I didn't know while studying.*

Usability

Two sub-themes were developed from the data obtained from students regarding the usability of ARUChatbot. Frequency data related to one of the usability themes, which is "liked features," are presented in Table 16.

Table 16. Frequency Table of Student Opinions on Liked Features

Liked Features	f
Providing instant feedback	7
Providing short and summarized information	6
Providing creative answers	3
Ability to make jokes, giving human-like answers	3
Ability to chat on topics outside of lesson	2

Examples of findings related to the "liked features" category are as follows:

E1: *...The short and clear responses, and the quickness of the answers were great.*

E2: *...I initially thought it was only related to this course, but I liked that it also answered questions about other courses.*

E3: *...When I asked a question, it would ask if I wanted to learn more details after the answer, or if I was interested in related topics. I really appreciated these suggestions.*

E4: *...When I first heard about it, I thought it would just provide standard responses like a search engine. However, it offered more creative answers.*

E5: *...It responded in different roles, such as acting as the course instructor. Additionally, it could make jokes, which was nice.*

E6: *...Providing instant, clear, and precise answers.*

E7: *...I liked its ability to provide immediate responses, make jokes, and summarize texts.*

E8: *...It felt like there was someone on the other side; chatting with it was enjoyable.*

E9: *...Compared to normal search engines, it was simpler and more explanatory. It didn't have too many sources, so it provided direct information without confusion or wasting time.*

Frequency data related to the "disliked features" category, another usability theme, are presented in Table 17.

Table 17. Frequency Table of Student Opinions on Disliked Features

Disliked Features	f
No visual or audio input-output, only text-based operation	9
Not saving message history, deleting messages when exiting the session	3
Not being easily accessible, no mobile application	2

According to the information presented in Table 17, all students (n=9) who participated in the interview and used ARUChatbot noted the absence of visual or audio input-output as a disliked feature. Some students (n=3) expressed their dissatisfaction with the deletion of their messages along with the session, indicating that they had to log in again every time the session closed. Additionally, 2 students mentioned that access to ARUChatbot was difficult and suggested that it would be more accessible if it were available as a mobile application.

Examples of findings related to the "disliked features" category are as follows:

E1: *...It did not continue from the previous question. Messages were deleted.*

E3: *...hen I logged out and logged back in or when my internet connection was lost, the messages were deleted, and I had to start over.*

E4: *...It could have been more accessible.*

E5: *...The lack of audio and visual support was unfavorable.*

E5: *...I did not like the absence of visual support.*

E7: *...I expected it to be like GPT, but it had fewer features compared to GPT, especially in terms of audio and visual support.*

E9: *...There could have been visual support; it would have been better if it were not just text based. Audio input would have made it easier.*

New Ideas

Two sub-themes related to the "new ideas" theme emerged from the interviews with students who used ARUChatbot.

(i) New ideas for design

Frequencies related to the student feedback on the "new design ideas" theme are presented in Table 18.

Table 18. Frequency Table of Student Opinions on New Ideas for Design

New Ideas for Design	f
Visual data input-output support	9
Stay logged in option without having to log in every time and message history not being deleted	3
Easy access in the form of a mobile application	2
Being able to chat with more human-like reactions	2
Having light/dark modes for day and night use	1
Having an entertainment module such as quiz and word game	1
Having a frequently asked questions section	1
Chatting with emojis more often	1

Here are some sample findings from the "new ideas for design" category based on student feedback:

O1: *"It could be converted into a mobile app. Messages could be sent to remind users. If it were integrated with the course, the lessons could be more enjoyable. For instance, if the teacher asked a question, everyone would research and try to answer it at that moment, increasing interaction."*

O3: *"It would have been better if sessions remained open and did not require re-entering email and password each time. Additionally, there could be a FAQ section or a help box."*

O4: *"It could be made more accessible."*

O5: *"It would be better if there was voice and image support. Sometimes I didn't want to type, but since there was no voice input, I couldn't. ARUChatbot met my expectations, but if it used natural language, I would use it more frequently."*

O7: *"There could have been an option for a dark or light mode in the design."*

O8: *"It would be better if it included visuals, not just text. The visual design could be improved, or there could be games like trivia or word games included."*

O9: *"If it saved the password, I wouldn't have to enter information each time and would use it more frequently."*

(ii) New ideas for the implementation process

Here are some examples of student feedback categorized under new ideas for implementing ARUChatbot:

- Integration with Classes and Permission for In-Class Use (n=5)
- Increase in Assignment Frequency (n=2)

Sample Findings:

O1: *"If we could get immediate answers to questions that come up while learning information in class, the knowledge might be more durable."*

O3: *"If the frequency of assignments increased, it would change how often I use it. I would use it for every assignment."*

O6: *"If this chatbot were used across the entire university and integrated, I would use it more frequently. If I could ask questions during class, it would make what I learn more lasting."*

O8: *"If we had used it in class as well, we would use it more often. We would understand what's going on better."*

Reasons for Non-use

Here are the reasons for not using ARUChatbot as reported by some students from the experimental group, along with examples of their feedback:

- Weak Initial Introduction and Lack of Understanding of Its Use (n=5)
- No Follow-up Training on How to Use After the Initial Introduction (n=2)
- Lack of Support at the Time of Need and Time Constraints (n=*2)*

Sample Findings:

O10: *"I registered for ARUChatbot but didn't understand how to use it and didn't want to ask you again. If there had been help available, I would have learned how to use it and would have used it."*

O12: *"I didn't know how to use it. If I had known you, we could have communicated, and I would have used it. If you had shown me how to use it face-to-face, I would have used it."*

O13: *"I was not aware of ARUChatbot. This is the first time I'm hearing about it."*

O15: *"I don't understand much about computers, so I didn't use it because I didn't know how. If the classes had been face-to-face instead of online, and if I had known you, we could have communicated, and I would have used it."*

DISCUSSIONS AND CONCLUSION

RQ-1: What Is the Effect of Instructional Chatbot Usage on Academic Achievement?

The results related to this research question indicate that the use of instructional chatbots positively impacts students' academic performance. The findings reveal that students in the experimental group achieved higher academic success compared to those in the control group. This outcome may be attributed to the increased time that students in the experimental group spent on course-related activities outside of class by using the chatbot. Specifically, these students engaged more with course activities during their out-of-class time through the chatbot to learn or understand course topics better. The strong and significant relationship detected between the number of messages sent by students in the experimental group and their academic success further supports this finding. Similar results have been reported in other studies examining the impact of instructional chatbot use on academic performance, which have also indicated that chatbot usage positively affects academic success (Chang et al., 2022; Chen et al., 2020; Deveci et al., 2021; Essel et al., 2022; Lee et al., 2022; Li, 2023; Song & Kim, 2021). Instructional chatbots may contribute to better understanding and higher academic achievement by enriching the learning experience, providing instant feedback, and personalizing the learning process. Instant feedback supports learning and has positive effects on learners' academic success (Grotz, 2016; Karadeniz, 2009; Nagy et al., 2006; Narciss & Huth, 2006; Tscholl et al., 2022; Wu et al., 2012). Personalized learning environments offer students the opportunity to integrate their out-of-school interests, values, and knowledge directly into the learning process. This can help students better understand challenging topics by providing content that matches their individual learning speeds and styles, thus enhancing their overall academic performance (Hwang et al., 2012; Roberts et al., 2017; Tseng et al., 2008; Walkington & Bernacki, 2014). Other studies examining the impact of instructional chatbots on education have also found that these chatbots enrich the learning experience (Ferrandez et al., 2024; Kerly et al., 2006; Tamayo et al., 2020), provide personalized and instant feedback (Ait Baha et al., 2023; Chuang et al., 2023; Hobert, 2021; Kohnke, 2021; Lin & Ye, 2023; Saez-Fernandez et al., 2019). Instructional chatbots may have facilitated greater participation in course-related processes outside of class by offering more interaction opportunities. Studies in the literature have reported similar findings (Chuang et al., 2023; Deveci et al., 2021; Mokmin et al., 2021; Pereira et al., 2019; Tamayo et al., 2020; Zhang et al., 2022). This increased participation may have indirectly led to enhanced student motivation and, consequently, improved academic success. Studies that indicate instructional chatbots increase students' motivation support this outcome (Ait Baha et al., 2023; Deveci et al., 2021; Hobert, 2021; Lee et al., 2022; Lin & Ye, 2023). Interaction during the learning process is a significant variable affecting students' engagement, satisfaction, motivation, and academic success (Anderson, 2006; Bernard et al., 2009; Moore & Kearsley, 1996; Wang et al., 2014). In particular, interaction in online education based on distance learning enhances students' course participation (Matei & Ball-Rokeach, 2001). Motivation, a key component of the learning process, also enhances students' learning performance and success (Keller & Suzuki, 2004).

RQ-2: What Types of Learning Activities Do Students Write Messages About? What Is the Distribution of Messages Written by Students Concerning Learning Activities?

The various learning activities conducted using the instructional chatbot have been identified as information retrieval, assignment completion, course preparation, and reinforcement. Activities such as taking class notes or obtaining information from online sources can contribute to students' academic success. Kiewra et al. (1991) demonstrated that note-taking skills help students focus better on the class and understand course materials more thoroughly, thereby enhancing their academic achievement. Means et al. (2013) indicated that online resources provide students with broader access and offer more flexible participation in learning. The process of completing assignments can enhance students' abilities to process, analyze, and synthesize learning materials, which can lead to improved academic performance (Trautwein & Koller, 2003). Cooper et al. (2006) found that completing assignments improves academic success, highlighting those assignments provide opportunities for students to apply and reinforce what they have learned. Zimmerman and Kitsantas (2005) stated that completing assignments helps develop students' self-regulation skills, which contribute to academic success. Assignments aid in time management, goal setting, and developing learning strategies. Using instructional chatbots to come prepared for class can enhance classroom interactions and help students learn course materials more effectively. The impact of being prepared for class on academic success is supported by studies. For instance, Kiewra (2002) emphasized that students who review course materials beforehand show more active participation during class, which increases their academic success. Prepared students engage more in class interactions and process information more deeply, leading to higher academic performance. Hashim (2023) suggested that being prepared for class fosters active learning and increases engagement in the learning process, noting that the time spent preparing is similar to time spent studying for exams. Regular review of course materials using instructional chatbots can enhance information retention and improve exam performance. Regular review and reinforcement of knowledge are critical for ensuring learning retention. Ebbinghaus's (2013) forgetting curve studies show that information is forgotten over time and that regular review helps maintain information in long-term memory. Reinforcement aids in transferring learned information to long-term memory and improving exam performance. Cepeda et al. (2006) demonstrated the positive effects of spaced repetition on long-term learning and recall. Based on these findings, it can be said that integrating instructional chatbots into personalized learning environments allows students to learn at their own pace and manage their learning processes, thereby contributing to their academic success. Additionally, instructional chatbots have the potential to enhance flexibility in students' individual study processes, making them an alternative tool for personalized learning in both face-to-face and distance education settings. Research shows that instructional chatbots provide personalized teaching support and positively impact academic achievement (Ait Baha et al., 2023; Hobert, 2021; Kohnke, 2021; Lin & Ye, 2023; Saez-Fernandez et al., 2019).

RQ-3: What Are the Students' Tendencies Towards Using Instructional Chatbots?

In this research question, the variation in the use of instructional chatbots over time, particularly in relation to exam weeks, has been identified. It has been observed that there is a decreasing trend in the use of instructional chatbots after the initial week. This trend may indicate that students began using the instructional chatbot less over time. Possible reasons for this decline could include students having less need for the chatbot after obtaining the information they required, or the assistance provided by the chatbot being less beneficial than anticipated. In other words, students might have reduced their use of the chatbot either because their needs were met or because they found the assistance offered by the chatbot insufficiently helpful. Additionally, this decrease might be attributed to a reduction in learning effort in activities that do not involve preparation for assessments. Students' engagement and persistence in learning processes are dependent on their academic motivation levels (Artino & Stephens, 2009; Inci, 2019). Moreover, since this technology is not yet widely used in educational settings, there may also be a novelty effect where students used the chatbot more frequently during the first week due to curiosity. The introduction of recent technologies in education often captures students' attention and interest initially. This is generally due to the novelty and uniqueness of the technology. When students encounter a tool or method they have not experienced before, it can motivate them and increase their participation in the learning process. However, as the use of this technology becomes more routine over time, students' level of engagement with it may decrease. Kay and Lauricella (2011) observed that when recent technologies such as iPads are introduced to students, there is initially high interest and motivation, but this interest tends to wane over time. The effective use of technology and a lack of pedagogical integration can reduce student motivation and interest, leading to a decrease in the impact of the technology on learning (Muir-Herzig, 2004).

Another finding related to exam weeks indicates that the use of instructional chatbots peaks during this period. This suggests that students require more help during exams and use instructional chatbots more actively at this time. Exam periods are intense and stressful times for students. They need quick and easy access to information while preparing for exams (Ilaslan, 2023). Therefore, tools like instructional chatbots that provide instant feedback and support may be used more frequently during these times. Additionally, the increased use of instructional chatbots during exam periods and assignment deadlines may reflect a focus on exam-oriented educational processes. The intensive use of chatbots during these periods could also be due to their role as a resource for course content. Findings showing increased chatbot usage as assignment deadlines approach suggest that instructional chatbots are a significant support tool in education. They can positively affect academic success by addressing students' needs for solving questions, obtaining information, and clearing last-minute doubts. The increase in chatbot use as assignment deadlines approach may point to issues such as time management and procrastination. Students with lower self-regulation may tend to delay their assignments (Wolters & Hussain, 2015). Students driven by external motivation might postpone tasks as deadlines approach (Senecal et al., 1995). Students may require additional support to complete assignments at the last minute. The immediate availability of chatbots in such situations can help students manage these challenging periods better. Furthermore, students' increased desire to obtain more information as assignment deadlines approach signifies active support for their learning processes. Ariely and Wertenbroch (2002) noted that when students are given mandatory deadlines, they generally tend to adhere to these deadlines and show improved performance. Chatbots could potentially encourage deeper learning by providing additional information and resources on assignment topics.

To summarize the results related to this research question, it is evident that the use of instructional chatbots can vary over time and may be in higher demand during specific periods, such as intense times like exam weeks or as assignment deadlines approach. This information can be crucial for planning and managing the use of instructional chatbots, particularly in terms of adjusting the content and services of the chatbot to meet students' needs. Additionally, it can provide guidance on how to optimize instructional chatbots and similar tools to enhance students' academic success.

RQ-4: What Are Students' Experiences with Using Chatbots?

When evaluating the usability of instructional chatbots, students prioritize features such as immediate responses, providing concise information, and the ability to converse on non-academic topics. However, there are also criticisms of the chatbot, particularly regarding its text-based functionality and the lack of certain features. In response to these criticisms, students have suggested various improvements for the design and operation of instructional chatbots. Notable suggestions include support for visual or audio input/output, easy access via mobile applications, and the option to keep sessions open. Chiu et al. (2023) examined the impact of AI-based chatbots on student motivation and emphasized that such additional features significantly impact learning processes. In the study conducted by Li (2023), they observed the performance of students in their projects with the integration of a ChatGPT-based flipped learning approach. They revealed how such additional features contribute to success. Adding visual and auditory elements to the chatbot can increase student satisfaction and affect the frequency of use. Additionally, recommendations for allowing in-class use of instructional chatbots and increasing the frequency of assignments are important. These results indicate that adding or improving specific features and functionalities could enhance the user experience with instructional chatbots. Considering such suggestions could help make instructional chatbots more effective and user-friendly. Literature suggests that instructional chatbots offered as mobile applications are more accessible and that easy access affects the frequency of chatbot use (Chang et al., 2022; Kohnke, 2021). Moreover, Ait Baha et al. (2023) highlight that integrating instructional chatbots into smart classrooms can enhance students' interest in classes and improve academic success.

Some students in the experimental group chose not to use the instructional chatbot. Findings indicate that there are several reasons for this preference. Factors such as insufficient introduction to the instructional chatbot and lack of adequate training on its use might have contributed to students' reluctance. In such cases, it is important to effectively promote the potential benefits and usage methods of instructional chatbots to students. Additionally, issues like not receiving help when needed or time constraints have also been noted as factors affecting the decision to avoid using the chatbot. Therefore, making the chatbot's use more flexible and accessible is crucial to providing timely and effective support to students. These results suggest that to increase the use of instructional chatbots and reach a broader student audience, improvements in promotion, user-friendly training, and support services are necessary. By addressing these areas, the full potential of instructional chatbots can be realized, providing greater benefits to students.

RECOMMENDATIONS

Studies conducted in different countries and cultures can enhance the generalizability of the findings. In particular, multicultural contexts help us understand how the effects of instructional chatbots on learning vary across different groups. The impact of diverse cultural structures and educational settings on chatbot usage can indicate whether ARUChatbot would be similarly effective among student groups in other countries. Additionally, while observing the effects of instructional chatbots, attention can be given to the use of different educational materials and technological tools. Analyses conducted with various study groups can reveal how learning outcomes and the impact of technology differ. These approaches help us understand whether the chatbot is more effective merely as a tool or when integrated with a specific educational strategy.

Based on the results obtained from the study, the following recommendations are made for researchers who will investigate the design of instructional chatbots and for future studies exploring the effects or application aspects of instructional chatbots:

Recommendations for Researchers Who Will Conduct Research on Instructional Chatbot Design

Recommendations for researchers focusing on instructional chatbot design can aid in developing more user-friendly and effective chatbots by considering student feedback and user experiences. Below are some suggestions for researchers in this field:

i. Student Feedback Integration

Consideration should be given to students' suggestions and critiques regarding the design and functionality of instructional chatbots. Student feedback is critical for tailoring chatbots to user needs. Norman and Draper (1986) emphasize the importance of user-centered design, noting that integrating user feedback into the product development process enhances user satisfaction. Kay and LeSage (2009) have demonstrated that student feedback is crucial in the development of educational technologies, and such feedback significantly improves the effectiveness of these technologies.

ii. Visual and Auditory Input-Output Support

Providing visual or auditory input-output support in instructional chatbots can enhance student interaction and accessibility. Such features may also be beneficial for students with different learning styles. Mayer (2002) supports the multimodal learning theory, demonstrating that learning supported by visual and auditory information is more effective. Lazar et al. (2007) emphasizes the importance of technology accessibility, noting that visual and auditory supports improve the user experience.

iii. Mobile Application and Easy Access

Designing instructional chatbots as mobile applications and ensuring easy access can enable students to use these technologies more frequently and effectively. Features such as session persistence can enhance the user experience. Crompton (2013) notes that mobile learning (m-learning) offers flexibility and accessibility in education, allowing students to access learning materials anytime and anywhere. Nielsen (1993) underscores the importance of usability principles in technology design, stating that easy access and user-friendly interfaces improve student satisfaction.

iv. Assessment in Various Learning Contexts

Studies should be conducted to evaluate the effectiveness of instructional chatbots in different learning contexts, such as classroom settings, online education, and blended learning environments. These evaluations can help understand how chatbots can be used more effectively across various educational settings. Lave and Wenger (1991) argue that learning is contextual and that learning processes occur differently across various contexts. This highlights the importance of investigating how chatbots can be effective in different educational environments. Dede (2005) suggests that the adaptability of educational technologies—specifically, their ability to be tailored to students' needs and contexts—enhances learning effectiveness.

v. Ensuring User-Friendly and Effective Design of Instructional Chatbots

Further research should be conducted to ensure that instructional chatbots are user-friendly and effective. These studies should aim to understand how chatbots can support learning processes and enhance student satisfaction. Shneiderman (1998) emphasizes the importance of human-computer interaction, demonstrating that user-friendly designs improve user experience and satisfaction. Laurillard (2012) notes that the effective design of educational technologies contributes to learning processes and enhances student success.

Recommendations for Researchers Who Will Conduct Research on Instructional Chatbot Applications and Their Effects

Recommendations for researchers investigating instructional chatbot applications could encourage the effective and widespread use of these tools in education. These recommendations may help develop strategies for the utilization, promotion, and implementation of chatbots by considering student feedback and findings from the literature.

i. Promoting and Information Studies Emphasizing the Benefits of instructional Chatbot

Highlighting findings that demonstrate how instructional chatbots contribute to students' learning processes could support the widespread adoption of these tools. Raising awareness among students and teachers about the benefits of chatbots can increase the adoption of these technologies. Davis (1989) emphasized that perceived usefulness and ease of use are critical in the adoption of technologies. Informational campaigns that highlight the benefits of chatbots can facilitate the acceptance of these technologies by users. Rogers (2003) noted that informing and promoting innovations are crucial for their adoption. Focused efforts on such initiatives are essential for the broader integration of AI-supported tools in education.

ii. Push Processes and Reminders

Given that students use chatbots more intensively on assignment due dates and during exam weeks, it is important to ensure that chatbots remain active throughout the entire learning process. Implementing push notifications to provide students with topic reminders or suggestions can help continuously support their learning activities. Cepeda et al. (2006) demonstrated the positive effects of spaced repetition on learning. Push notifications can assist students in regularly reviewing learning materials. Zimmerman (2002) highlighted that self-regulated learning improves students' academic achievements. Supporting this process through chatbots can enable students to manage their learning activities more effectively.

iii. User Training and Support Services

Insufficient training for students on how to use instructional chatbots or the lack of help support when needed can negatively impact the utilization of these technologies. Therefore, it is essential to prioritize user training and support services. Gagne et al. (2005) indicated that effective training design supports learning wwthe importance of effective support services in increasing user satisfaction with information systems. Similar support should be provided for chatbot usage.

iv. Encouraging Student Use of Instructional Chatbots

To increase the use of instructional chatbots, promotional activities targeted at students can be organized. Additionally, informative materials detailing how chatbots contribute to academic success can be developed. Deterding et al. (2011) demonstrated that gamification enhances user motivation and engagement. Incentive campaigns directed at students could boost the utilization of chatbots. Fullan (2001) emphasized the importance of informative materials and dissemination efforts in the adoption of innovations in education. Materials that explain the benefits of chatbots can support the adoption of these technologies.

Authors' Note: This study is derived from the doctoral thesis titled "Instructional Chatbot Practice in Distance Education and Investigation of Its Role in the Learning Process," conducted at the Department of Computer and Instructional Technologies Education, Institute of Educational Sciences, Ataturk University.

Acknowledgements: We would like to acknowledge that we have utilized ChatGPT-4 and Grammarly for the purpose of proofreading and enhancing the clarity of our expressions, as well as correcting any typographical errors in this article.

BIODATA and CONTACT ADDRESSES of AUTHORS

Dr. Ahmet Yusuf CEVHER is an Assistant Professor of Computer Engineering at the Faculty of Engineering, Ardahan University. He received his Ph.D. in Computer Education and Instructional Technology from Ataturk University in 2024, with a dissertation on instructional chatbot applications in distance education and their role in the learning process. His main research interests include educational technologies, artificial intelligence in education, human-computer interaction, open and distance learning, and instructional chatbots. He has published articles in internationally indexed journals, contributed to international book chapters, and presented papers at national and international conferences. In addition, he has taken part in and coordinated various research projects supported by TUBITAK, the European Union, and higher education institutions, particularly focusing on artificial intelligence, educational technologies, and digital transformation.

Ahmet Yusuf CEVHER
Department of Computer Engineering, Faculty of Engineering
Address: Ardahan University, 75002, Ardahan, Turkiye
Phone: +90 5397843160
E-mail: ahmetyusufcevher@ardahan.edu.tr

Dr. Serkan YILDIRIM is an Associate Professor of Computer Education and Instructional Technology at Ataturk University, Kazim Karabekir Faculty of Education. He received his Ph.D. from Gazi University in 2012 with a dissertation focusing on the comparison of virtual world and web-based learning environments in terms of students' academic achievement, motivation, and social presence. His research interests include open and distance learning, digital learning environments, virtual and augmented reality in education, learning analytics, and educational technologies. He has published articles in national and international journals, book chapters, and conference proceedings, as well as serving as an editor and author of academic books in the field of digital literacy, infographics, and online learning. He has also coordinated and contributed to several national and institutional research projects, particularly on mobile learning, artificial intelligence in education, and assessment system.

Serkan YILDIRIM
Department of Computer Education and Educational Technologies, Kazim Karabekir Faculty of Education
Address: Ataturk University, 25240, Erzurum, Turkiye
Phone: +90 5053131329
E-mail: serkanyildirim@atauni.edu.tr

REFERENCES

Adamopoulou, E., & Moussiades, L. (2020). Chatbots: History, technology, and applications. *Machine Learning with Applications*, 2, 100006. https://doi.org/10.1016/j.mlwa.2020.100006

Adiguzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology, 15*(3). ep429. https://doi.org/10.30935/cedtech/13152

Ahmad, F. K. (2015). Use of assistive technology in inclusive education: making room for diverse learning needs. *Transcience*, *6*(2), 62-77.

Ahn, J., Lee, J., & Son, M. (2024). ChatGPT in ELT: disruptor? Or well-trained teaching assistant?. *ELT Journal, 78*(3), 345-355.

Ait Baha, T., El Hajji, M., Es-Saady, Y., & Fadili, H. (2023). The impact of educational chatbot on student learning experience. *Education and Information Technologies*, 1-24. https://doi.org/10.1007/s10639-023-12166-w

Altwaijry, N., Ibrahim, A., Binsuwaidan, R., Alnajjar, L. I., Alsfouk, B. A., & Almutairi, R. (2021). Distance education during COVID-19 pandemic: A college of pharmacy experience. *Risk Management and Healthcare Policy*, 2099-2110. https://doi.org/10.2147/RMHP.S308998

Anderson, T. (2006). Interaction in learning and teaching on the educational semantic web. *Interactions in Online Education* (pp. 141-155). Routledge.

Ariely, D., & Wertenbroch, K. (2002). Procrastination, deadlines, and performance: Self-control by precommitment. *Psychological science, 13*(3), 219-224.

Arslan, G. B., Kizilay, E., & Hamalosmanoglu, M. (2022). Egitimde teknoloji entegrasyonu ile ilgili Turkiye'de yapilan calismalarin incelenmesi. *Anadolu Universitesi Egitim Fakultesi Dergisi*, *6*(1), 39-55. https://doi.org/10.34056/aujef.976627

Artino, A. R., & Stephens, J. M. (2009). Academic motivation and self-regulation: A comparative analysis of undergraduate and graduate students learning online. *Internet and Higher Education*, *12*(3), 146-151. https://doi.org/10.1016/j.iheduc.2009.02.001

Aydin, H., Mutlu, M. E., Karadag, N., Usta, I., Aydin, S., Firat, M., & Ozdamar Keskin, N. (2015). Temel kavramlar, sozcuk islemler, sunum teknolojileri, hesap tablolari, internet teknolojileri. Yuzer, T., V. & Okur, M., R. (Eds.), *Temel Bilgi Teknolojileri I* (ss. 1-124). Anadolu Universitesi Yayinlari. https://ets.anadolu.edu.tr/storage/nfs/BIL101U/ebook/BIL101U-15V1S1-8-0-1-SV1-ebook.pdf

Bai, S., Lo, C. K., & Yang, C. (2024). Enhancing instructional design learning: a comparative study of scaffolding by a 5E instructional model-informed artificial intelligence chatbot and a human teacher. *Interactive Learning Environments,* 1-20.

Balaceanu, D. (2020). The origins of modern chatbots. https://www.druidai.com/blog/how-chatbots-were-invented-origins

Balyer, A., & Oz, O. (2018). Academicians' views on digital transformation in education. *International Online Journal of Education and Teaching*, *5*(4), 809-830.

Barnard, L., Lan, W. Y., To, Y. M., Paton, V. O., & Lai, S. L. (2009). Measuring self-regulation in online and blended learning environments. *The Internet and Higher Education, 12*(1), 1-6. https://doi.org/10.1016/j.iheduc.2008.10.005

Buyukozturk, S., Kilic Cakmak, E., Akgun, O. E., Karadeniz, S., & Demirel, F. (2008). *Bilimsel arastirma yontemleri* (17. baski). Pegem Akademi.

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132*(3), 354. https://doi.org/10.1037/0033-2909.132.3.354

Chandrasekar, V. (2023). What is the history of Google Assistant?. https://www.quora.com/What-is-the-history-of-Google-Assistant

Chang, C. Y., Hwang, G. J., & Gau, M. L. (2022). Promoting students' learning achievement and self-efficacy: A mobile chatbot approach for nursing training. *British Journal of Educational Technology*, *53*(1), 171-188. https://doi.org/10.1111/bjet.13158

Chang, Ching-Yi, Shu-Yu Kuo, & Gwo-Haur Hwang, (2022). Chatbot-facilitated nursing education. *Educational Technology & Society,* 25(1), 15-27.

Chen, C. Y., Yu, D., Wen, W., Yang, Y. M., Zhang, J., Zhou, M., ... & Yu, Z. (2018). Gunrock: Building a human-like social bot by leveraging large scale real user data. *Alexa prize proceedings*.

Chiu, T. K., Moorhouse, B. L., Chai, C. S., & Ismailov, M. (2023). Teacher support and student motivation to learn with Artificial Intelligence (AI) based chatbot. *Interactive Learning Environments*, 1-17. https://doi.org/10.1080/10494820.2023.2172044

Chuang, C. H., Lo, J. H., & Wu, Y. K. (2023). Integrating chatbot and augmented reality technology into biology learning during COVID-19. *Electronics, 12*(1), 222. https://doi.org/10.3390/electronics12010222

Clark, T. M. (2023). Investigating the use of an artificial Intelligence Chatbot with General Chemistry Exam Questions. *Journal of Chemical Education*, *100*(5), 1905-1916. https://doi.org/10.1021/acs.jchemed.3c00027

Cooper, H., Robinson, J. C., & Patall, E. A. (2006). Does homework improve academic achievement? A synthesis of research, 1987–2003. *Review of Educational Research*, *76*(1), 1-62. https://doi.org/10.3102/00346543076001001

Crawford, J., Allen, K. A., Pani, B., & Cowling, M. (2024). When artificial intelligence substitutes humans in higher education: the cost of loneliness, student success, and retention. *Studies in Higher Education,* 1-15. https://doi.org/10.1080/03075079.2024.2326956

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches* (4nd ed.). Sage.

Crompton, H. (2013). A historical overview of m-learning: Toward learner-centered education. In *Handbook of mobile learning* (pp. 3-14). Routledge.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319-340. https://doi.org/10.2307/249008

Dede, C. (2005). Planning for neomillennial learning styles. *Educause Quarterly*, *28*(1), 7-12.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011, September). From game design elements to gamefulness: defining" gamification". In *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments* (pp. 9-15). https://doi.org/10.1145/2181037.2181040

Deveci Topal, A., Dilek Eren, C., & Kolburan Gecer, A. (2021). Chatbot application in a 5th grade science course. *Education and Information Technologies*, *26*(5), 6241-6265. https://doi.org/10.1007/s10639-021-10627-8

Ebbinghaus, H. (2013). Memory: A contribution to experimental psychology. *Annals of Neurosciences, 20*(4), 155. https://doi.org/10.5214/ans.0972.7531.200408

Edisherashvili, N., Saks, K., Pedaste, M., & Leijen, A. (2022). Supporting self-regulated learning in distance learning contexts at higher education level: systematic literature review. *Frontiers in Psychology*, *12*, 792422. https://doi.org/10.3389/fpsyg.2021.792422

Essel, H. B., Vlachopoulos, D., Tachie-Menson, A., Johnson, E. E., & Baah, P. K. (2022). The impact of a virtual teaching assistant (chatbot) on students' learning in Ghanaian higher education. *International Journal of Educational Technology in Higher Education, 19(*1), 57. https://doi.org/10.1186/s41239-022-00362-6

Fadhil, A. (2018). Beyond patient monitoring: Conversational agents role in telemedicine & healthcare support for home-living elderly individuals. *arXiv preprint arXiv:1803.06000.*

Ferrandez, A., Lavigne-Cervan, R., Peral, J., Navarro-Soria, I., Lloret, A., Gil, D., & Rocamora, C. (2024). CuentosIE: can a chatbot about "tales with a message" help to teach emotional intelligence?. *PeerJ Computer Science, 10*, e1866. https://doi.org/10.7717/peerj-cs.1866

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Sage.

Fullan, M. (2015). *The new meaning of educational change*. Teachers College.

Gagne, R. M., Wager, W. W., Golas, K. C., Keller, J. M., & Russell, J. D. (2005). *Principles of instructional design*.

Gastwirth, J. L., Gel, Y. R., & Miao, W. (2009). The impact of Levene's test of equality of variances on statistical theory and practice. https://doi.org/10.1214/09-STS301

Grotz, N. (2016, April). Improving programming skills using computer based feedback and peer group competition. *In 2016 IEEE Global Engineering Education Conference* (EDUCON) (pp. 585-591). [Symposium]. IEEE. https://doi.org/10.1109/EDUCON.2016.7474610

Gunawardena, C. N., & McIsaac, M. S. (2013). Distance education. In *Handbook of research on educational communications and technology* (pp. 361-401). Routledge.

Hashim, H. (2023). Early class self-preparation: a step ahead for a better learning experience. *Buletin APB Edisi 11*.

Heiser, J. F., Colby, K. M., Faught, W. S., & Parkison, R. C. (1979). Can psychiatristsdistinguish a computer simulation of paranoia from the real thing?: The limitationsof turing-like tests as measures of the adequacy of simulations. *Journal of Psychiatric Research*, *15*(3), 149–162. https://doi.org/10.1016/0022-3956(79)90008-6

Hobert, S. & Wolff, R.M. (2019, February). Say hello to your new automated tutor- a structured literature review on pedagogical conversational agents. 1*4th International Conference on Wirtschaftsinformatik* [Symposium]. Germany. https://core.ac.uk/download/pdf/301380749.pdf sayfasindan erisilmistir.

Hobert, S. (2021, November). Individualized learning patterns require individualized conversations–data-driven insights from the field on how chatbots instruct students in solving exercises. *In International Workshop on Chatbot Research and Design* (pp. 55-69). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-94890-0_4

Hochberg Y.& Tamhane, A.C. (1987). *Multiple comparison procedures.* John Wiley & Sons press.

Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning*, *38*(1), 237-257. https://doi.org/10.1111/jcal.12610

Humphry, T., & Fuller, A. L. (2023). Potential ChatGPT Use in Undergraduate Chemistry Laboratories. *Journal of Chemical Education*, *100*(4), 1434-1436. https://doi.org/10.1021/acs.jchemed.3c00006

Hwang, G. J., & Chang, C. Y. (2021). A review of opportunities and challenges of chatbots in education. *Interactive Learning Environments*. https://doi.org/10.1080/10494820.2021.1952615

Hwang, G. J., Sung, H. Y., Hung, C. M., Huang, I., & Tsai, C. C. (2012). Development of a personalized educational computer game based on students' learning styles. *Educational Technology Research and Development, 60*, 623-638. https://doi.org/10.1007/s11423-012-9241-x

Imran, M., & Almusharraf, N. (2023). Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature. *Contemporary Educational Technology*, *15*(4), ep464. https://doi.org/10.30935/cedtech/13605

Ilaslan, E. (2023). Yapay zekâ sohbet robotlari ve ChatGPT'nin hemsirelik egitiminde kullanilmasi. *Akdeniz Hemsirelik Dergisi*, *2*(2), 73-80. https://doi.org/10.59398/ahd.1330341

Inci, T. (2019). *Baglam temelli ogrenme ortami algisi, derse ilgi, derse katilim ve akademik gudulenme etkilesiminin ortaokul ogrencilerinin fen bilimleri basarisina etkisi.* [Yayimlanmamis Doktora Tezi]. Eskisehir Osmangazi Universitesi.

Jeon, J., & Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies*, 1-20. https://doi.org/10.1007/s10639-023-11834-1

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

Karadeniz, S. (2009). The impacts of paper, web and mobile based assessment on students' achievement and perceptions. *Scientific Research and Essay, 4*(10), 984-991.

Karakoc, Y. F., & Donmez, L. (2014). Olcek gelistirme calismalarinda temel ilkeler. *Tip Egitimi Dunyasi, 13*(40), 39-49. https://doi.org/10.25282/ted.228738

Kay, R. H., & LeSage, A. (2009). Examining the benefits and challenges of using audience response systems: A review of the literature. *Computers* & *Education*, *53*(3), 819-827. https://doi.org/10.1016/j.compedu.2009.05.001

Kay, R., & Lauricella, S. (2011). Exploring the benefits and challenges of using laptop computers in higher education classrooms: A formative analysis. *Canadian Journal of Learning and Technology/La revue canadienne de l'apprentissage et de la technologie, 37*(1). https://doi.org/10.21432/T2S598

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

Keller, J., & Suzuki, K. (2004). Learner motivation and e-learning design: A multinationally validated process. *Journal of educational Media*, *29*(3), 229-239. https://doi.org/10.1080/1358165042000283084

Kerly, A., Hall, P., & Bull, S. (2006, December). Bringing chatbots into education: Towards natural language negotiation of open learner models. In *International conference on innovative techniques and applications of artificial intelligence* (pp. 179-192). Springer London. https://doi.org/10.1007/978-1-84628-666-7_14

Kiewra, K. A. (2002). How classroom teachers can help students learn and teach them how to learn. *Theory into Practice, 41*(2), 71-80. https://doi.org/10.1207/s15430421tip4102_3

Kiewra, K. A., DuBois, N. F., Christian, D., McShane, A., Meyerhoffer, M., & Roskelley, D. (1991). Note-taking functions and techniques. *Journal of Educational Psychology, 83*(2), 240. https://doi.org/10.1037/0022-0663.83.2.240

Kim, T. K., & Park, J. H. (2019). More about the basic assumptions of t-test: normality and sample size. Korean Journal of Anesthesiology, 72(4), 331-335. https://doi.org/10.4097/kja.d.18.00292

Kohnke, L. (2021, November). Chatbots for Learning: A Facebook Messenger 'Bot'. In *International Symposium on Emerging Technologies for Education* (pp. 373-377). Springer International Publishing. https://doi.org/10.1007/978-3-030-92836-0_33

Kohnke, L. (2023). L2 learners' perceptions of a chatbot as a potential independent language learning tool. *International Journal of Mobile Learning and Organisation*, *17*(1-2), 214-226. https://doi.org/10.1504/IJMLO.2023.128339

Lai, J. W. (2024). Adapting Self-Regulated Learning in an Age of Generative Artificial Intelligence Chatbots. *Future Internet, 16*(6), 218.

Laurillard, D. (2013). *Teaching as a design science: Building pedagogical patterns for learning and technology*. Routledge.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.

Lawshe C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, *28*(4), 563-575. https://doi.org/10.1111/j.1744-6570.1975.tb01393.x

Lazar, J., Feng, J. H., & Hochheiser, H. (2017). *Research methods in human-computer interaction*. Morgan Kaufmann.

Lee, Y. F., Hwang, G. J., & Chen, P. Y. (2022). Impacts of an AI-based cha bot on college students' after-class review, academic performance, self-efficacy, learning attitude, and motivation. *Educational Technology Research and Development, 70*(5), 1843-1865. https://doi.org/10.1007/s11423-022-10142-8

Li, H. (2023). Effects of a ChatGPT-based flipped learning guiding approach on learners' courseware project performances and perceptions. *Australasian Journal of Educational Technology, 39*(5), 40-58. https://doi.org/10.14742/ajet.8923

Lin, Y. T., & Ye, J. H. (2023). Development of an educational chatbot system for enhancing students' biology learning performance. *Journal of Internet Technology, 24*(2), 275-281. https://doi.org/10.53106/160792642023032402006

Liu, I., Liu, F., Xiao, Y., Huang, Y., Wu, S., & Ni, S. (2024). Investigating the Key Success Factors of Chatbot-Based Positive Psychology Intervention with Retrieval-and Generative Pre-Trained Transformer (GPT)-Based Chatbots. *International Journal of Human–Computer Interaction*, 1-12. https://doi.org/10.1080/10447318.2023.2300015

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

Markowitz, D. M. (2024). Can generative AI infer thinking style from language? Evaluating the utility of AI as a psychological text analysis tool. *Behavior Research Methods*, 1-12. https://doi.org/10.3758/s13428-024-02344-0

Matei, S., & Ball-Rokeach, S. J. (2001). Real and virtual social ties: Connections in the everyday lives of seven ethnic neighborhoods. *American Behavioral Scientist, 45*(3), 550-564. https://doi.org/10.1177/00027642010045003012

Mayer, R. E. (2002). Multimedia learning. In *Psychology of learning and motivation* (Vol. 41, pp. 85-139). Academic Press.

Means, B., Toyama, Y., Murphy, R., & Baki, M. (2013). The effectiveness of online and blended learning: A meta-analysis of the empirical literature. *Teachers college record, 115*(3), 1-47. https://doi.org/10.1177/016146811311500307

Mohamed, A. M. (2023). Exploring the potential of an AI-based Chatbot (ChatGPT) in enhancing English as a Foreign Language (EFL) teaching: perceptions of EFL Faculty Members. *Education and Information Technologies*, 1-23. https://doi.org/10.1007/s10639-023-11917-z

Mokmin, N. A. M., & Ibrahim, N. A. (2021). The evaluation of chatbot as a tool for health literacy education among undergraduate students. *Education and Information Technologies*, *26*(5), 6033-6049. https://doi.org/10.1007/s10639-021-10542-y

Molnar, G., & Szuts, Z. (2018, September). The role of chatbots in formal education. In *2018 IEEE 16th International Symposium on Intelligent Systems and Informatics (SISY)* (pp.

Moore, M. G., & Kearsley, G. G. (1996). *Distance education: A system view*. Wadsworth.

Muir-Herzig, R. G. (2004). Technology and its impact in the classroom. *Computers & Education, 42*(2), 111-131. https://doi.org/10.1016/S0360-1315(03)00067-8

Muniandy, J., & Selvanathan, M. (2024). ChatGPT, a partnering tool to improve ESL learners' speaking skills: Case study in a Public University, Malaysia. *Teaching Public Administration*, 01447394241230152. https://doi.org/10.1177/01447394241230152

Nagy, T. F., Bauer, W., Raeburn, S. P., & Kashy, E. (2006, October). Work In Progress:" Explain It To Me". *In Proceedings. Frontiers in Education. 36th Annual Conference* (pp. 1-2). [Symposium]. IEEE. https://doi.org/10.1109/FIE.2006.322430

Narciss, S., & Huth, K. (2006). Fostering achievement and motivation with bug-related tutoring feedback in a computer-based training for written subtraction. *Learning and Instruction, 16*(4), 310-322. https://doi.org/10.1016/j.learninstruc.2006.07.003

Ng, D. T. K., Tan, C. W., & Leung, J. K. L. (2024). Empowering student self-regulated learning and science education through ChatGPT: A pioneering pilot study. *British Journal of Educational Technology*.

Nielsen, J. (1993). Usability engineering. Morgan Kaufmann Publishers In. *Inc.–1993*.

Nikolic, S., Daniel, S., Haque, R., Belkina, M., Hassan, G. M., Grundy, S., ... & Sandison, C. (2023). ChatGPT versus engineering education assessment: A multidisciplinary and multi-institutional benchmarking and analysis of this generative artificial intelligence tool to investigate assessment integrity. *European Journal of Engineering Education*, 1-56. https://doi.org/10.1080/03043797.2023.2248042

Nimavat, K. & Champaneria, T. (2017). Chatbots: An overview types, architecture, tools and future possibilities *International Journal for Scientific Research & Development, 5*(7), 1019-1024. https://www.researchgate.net/publication/320307269

Norman, D. A., & Draper, S. W. (1986). *User centered system design; new perspectives on human-computer interaction*. L. Erlbaum Associates Inc.. https://doi.org/10.1201/b15703

Nuruzzaman, M. & Hussain, O. K. (2018, October). *A survey on chatbot implementation in customer service industry through deep neural networks. IEEE 15th International Conference on e-Business Engineering*'de [Symposium]. Xi'an, China. https://doi.org/10.1109/ICEBE.2018.00019

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

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

Ozmen, H. (2014). Nicel arastirma yaklasimina dayali yontemler. M. Metin (Ed.), *Egitimde bilimsel arastirma yontemleri* icinde (ss. 47-75). Pegem Akademi.

Pendy, B. (2023). From traditional to tech-infused: the evolution of education. *BULLET: Jurnal Multidisiplin Ilmu*, *2*(3), 767-777.

Pereira, J., Fernandez-Raga, M., Osuna-Acedo, S., Roura-Redondo, M., Almazan-Lopez, O., & Buldon-Olalla, A. (2019). Promoting learners' voice productions using chatbots as a tool for improving the learning process in a MOOC. *Technology, knowledge and learning*, *24*(4), 545-565. https://doi.org/10.1007/s10758-019-09414-9

Perez, J. Q., Daradoumis, T., & Puig, J. M. M. (2020). Rediscovering the use of chatbots in education: A systematic literature review. *Computer Applications in Engineering Education*, *28*(6), 1549-1565. https://doi.org/10.1002/cae.22326

Polverini, G., & Gregorcic, B. (2024). How understanding large language models can inform the use of ChatGPT in physics education. *European Journal of Physics*, *45*(2), 025701. https://doi.org/10.1088/1361-6404/ad1420

Reznik, L. (2009). General principles and purposes of computational intelligence. *Systems Science and Cybernetics*, *3,* 198-222.

Roberts, L. D., Howell, J. A., & Seaman, K. (2017). Give me a customizable dashboard: Personalized learning analytics dashboards in higher education. *Technology, Knowledge and Learning, 22*, 317-333. https://doi.org/10.1007/s10758-017-9316-1

Rogers, E. M. (2003). Diffusion of innovations (5th ed.). Free.

Saez-Fernandez, M. D., Escobar, P., Marco-Such, M., & Candela, G. (2019). TUTOR-BOTS: a conversational chatbot for education. *In INTED2019 Proceedings* (pp. 8718-8722). IATED. https://doi.org/10.21125/inted.2019.2175

Senecal, C., Koestner, R., & Vallerand, R. J. (1995). Self-regulation and academic procrastination. *The journal of social psychology, 135*(5), 607-619. https://doi.org/10.1080/00224545.1995.9712234

Shneiderman, B., & Plaisant, C. (2010). *Designing the user interface: strategies for effective human-computer interaction*. Pearson Education India.

Shum, H. Y., He, X. D., & Li, D. (2018). From Eliza to XiaoIce: challenges and opportunities with social chatbots. *Frontiers of Information Technology* & *Electronic Engineering*, *19*(1), 10-26. https://doi.org/10.1631/FITEE.1700826

Smutny, P., & Schreiberova, P. (2020). Chatbots for learning: A review of educational chatbots for the Facebook Messenger. *Computers* & *Education*, *151*, 103862. https://doi.org/10.1016/j.compedu.2020.103862

Song, D., & Kim, D. (2021). Effects of self-regulation scaffolding on online participation and learning outcomes. *Journal of Research on Technology in Education, 53*(3), 249-263. https://doi.org/10.1080/15391523.2020.1767525

Sencan, H. (2005). Sosyal ve davranissal olcumlerde guvenilirlik ve gecerlilik.

https://books.google.com.tr/books?hl=tr&lr=&id=_AdLDwAAQBAJ&oi=fnd&pg=PA288&dq=G%C3%BCvenirlik,+ara%C5%9Ft%C4%B1rma+s%C3%BCrecinin+ve+kullan%C4%B1lan+%C3%B6l%C3%A7me+ara%C3%A7lar%C4%B1n%C4%B1n+istikrar%C4%B1n%C4%B1+yans%C4%B1t%C4%B1r+&ots=UMTm_5YTiq&sig=DlnG-fs2EhoXj_o-rSTjF0r1cpY&redir_esc=y#v=onepage&q&f=falsesayfasindan erisilmistir.

Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2013). *Using multivariate statistics* (Vol. 6, pp. 497-516). Pearson.

Tam, W., Huynh, T., Tang, A., Luong, S., Khatri, Y., & Zhou, W. (2023). Nursing education in the age of artificial intelligence powered Chatbots (AI-Chatbots): Are we ready yet?. *Nurse Education Today*, *129*, 105917. https://doi.org/10.1016/j.nedt.2023.105917

Tamayo, P. A., Herrero, A., Martin, J., Navarro, C., & Tranchez, J. M. (2020). Design of a chatbot as a distance learning assistant. *Open Praxis*, *12*(1), 145-153. https://doi.org/10.5944/openpraxis.12.1.1063

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

Trautwein, U., & Koller, O. (2003). The relationship between homework and achievement—still much of a mystery. *Educational Psychology Review*, *15*, 115-145. https://doi.org/10.1023/A:1023460414243

Tsang, R. (2023). Practical Applications of ChatGPT in Undergraduate Medical Education. *Journal of Medical Education and Curricular Development*, *10*, 23821205231178449. https://doi.org/10.1177/23821205231178449

Tscholl, M., Demir, F., & Kraner, C. (2022, July). Using Optical Head-Mounted Displays for instant, contextualized feedback. *In 2022 International Conference on Advanced Learning Technologies* (ICALT) (pp. 357-359). [Symposium]. IEEE. https://doi.org/10.1109/ICALT55010.2022.00111

Tseng, J. C., Chu, H. C., Hwang, G. J., & Tsai, C. C. (2008). Development of an adaptive learning system with two sources of personalization information. *Computers & Education, 51*(2), 776-786. https://doi.org/10.1016/j.compedu.2007.08.002

Tsivitanidou, O., & Ioannou, A. (2021). Envisioned pedagogical uses of chatbots in higher education and perceived benefits and challenges. In *International Conference on Human-Computer Interaction* (pp. 230-250). [Symposium]. Springer International Publishing. https://doi.org/10.1007/978-3-030-77943-6_15

Turing, A. M. (1950). Computing machinery and intelligence. *Mind, LIX 236*, 433–460. https://doi.org/10.1093/mind/LIX.236.433

Urhan, S., Gencaslan, O., & Dost, S. (2024). An argumentation experience regarding concepts of calculus with ChatGPT. *Interactive Learning Environments*, 1-26. https://doi.org/10.1080/10494820.2024.2308093

Verleger, M. & Pembridge, J. (2018, October). A pilot study integrating an AI-driven chatbot in an introductory programming course. *2018 IEEE Frontiers in Education Conference* [Symposium]. San Jose, USA. https://doi.org/10.1109/FIE.2018.8659282

Walkington, C., & Bernacki, M. L. (2014). Motivating students by "personalizing" learning around individual interests: A consideration of theory, design, and implementation issues. *In Motivational interventions* (Vol. 18, pp. 139-176). Emerald Group Publishing Limited. https://doi.org/10.1108/S0749-742320140000018004

Wallace, R. S. (2009). *The anatomy of ALICE* (pp. 181-210). Springer. https://doi.org/10.1007/978-1-4020-6710-5_13

Wang, Z., Chen, L., & Anderson, T. (2014). A framework for interaction and cognitive engagement in connectivist learning contexts. *The International Review of Research in Open and Distributed Learning, 15*(2). https://doi.org/10.19173/irrodl.v15i2.1709

Watts, F. M., & Finkenstaedt-Quinn, S. A. (2021). The current state of methods for establishing reliability in qualitative chemistry education research articles. *Chemistry Education Research and Practice, 22(3),* 565-578.

Weizenbaum, J. (1966). ELIZA-A computer program for the study of natural language communication between man and machine. *Communications of the ACM,* 9(1), 36-45. https://doi.org/10.1145/365153.365168

Winkler, R. & Sollner, M. (2018). Unleashing the potential of chatbots in education: A state-of-theart Analysis. *Academy of Management Annual Meeting (AOM)*, *1*, 15903. https://doi.org/10.5465/AMBPP.2018.15903abstract

Winkler, R., Sollner, M., & Leimeister, J. M. (2020, August). Improving Students' Problem-Solving Skills with Smart Personal Assistants. In *Annual Meeting of the Academy of Management (AOM).-Vancouver, Canada*. https://doi.org/10.5465/AMBPP.2020.11496abstract

Wollny, S., Schneider, J., Di Mitri, D., Weidlich, J., Rittberger, M., & Drachsler, H. (2021). Are we there yet?-A systematic literature review on chatbots in education. *Frontiers in Artificial Intelligence*, *4*, 654924. https://doi.org/10.3389/frai.2021.654924

Wolters, C. A., & Hussain, M. (2015). Investigating grit and its relations with college students' self-regulated learning and academic achievement. *Metacognition and Learning, 10*, 293-311. https://doi.org/10.1007/s11409-014-9128-9

Wu, P. H., Hwang, G. J., Milrad, M., Ke, H. R., & Huang, Y. M. (2012). An innovative concept map approach for improving students' learning performance with an instant feedback mechanism. *British Journal of Educational Technology, 43*(2), 217-232. https://doi.org/10.1111/j.1467-8535.2010.01167.x

Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies*, 1-25. https://doi.org/10.1007/s10639-023-11742-4

Yildirim, A., & Simsek, H. (1999). *Sosyal bilimlerde nitel arastirma yontemleri* (9. baski). Seckin.

Yin, J., Goh, T. T., & Hu, Y. (2024). Interactions with educational chatbots: the impact of induced emotions and students' learning motivation. *International Journal of Educational Technology in Higher Education, 21*(1), 47.

Yusuf, H., Money, A., & Daylamani-Zad, D. (2025). Pedagogical AI conversational agents in higher education: a conceptual framework and survey of the state of the art. *Educational technology research and development*, 1-60.

Zadorozhnyy, A., & Lai, W. Y. W. (2023). ChatGPT and L2 Written Communication: A Game-Changer or Just Another Tool?. *Languages*, *9*(1), 5. https://doi.org/10.3390/languages9010005

Zaikin, O., Tadeusiewicz, R., Rożewski, P., Kofoed, L. B., Malinowska, M., & Zyławski, A. (2016). Teachers' and students' motivation model as a strategy for open distance learning processes. *Bulletin Of The Polish Academy Of Sciences Technical Sciences*, 943-955. https://doi.org/10.1515/bpasts-2016-0103

Zhang, C., Li, G., Hashimoto, H., & Zhang, Z. (2022, March). Digital transformation (dx) for skill learners: the design methodology and implementation of educational chatbot using knowledge connection and emotional expression. *In 2022 IEEE Global Engineering Education Conference (EDUCON)*, (pp. 998-1003). [Symposium]. IEEE. https://doi.org/10.1109/EDUCON52537.2022.9766384

Zimmerman, B. J., & Kitsantas, A. (2005). Homework practices and academic achievement: The mediating role of self-efficacy and perceived responsibility beliefs. *Contemporary educational psychology, 30*(4), 397-417. https://doi.org/10.1016/j.cedpsych.2005.05.003

Zybaczynska, J., Norris, M., Modi, S., Brennan, J., Jhaveri, P., Craig, T. J., & Al-Shaikhly, T. (2024). Artificial intelligence–generated scientific literature: A critical appraisal. *The Journal of Allergy and Clinical Immunology: In Practice*, *12*(1), 106-110. https://doi.org/10.1016/j.jaip.2023.10.010