

# THE IMPACT OF ARTIFICIAL INTELLIGENCE CHATBOTS ON STUDENT LEARNING: A QUASI-EXPERIMENTAL ANALYSIS OF LEARNING OUTCOME AND ENGAGEMENT

Stella Eteng-Uket\*, University of Port Harcourt  
Ebere Ezeoguine, University of Port Harcourt

---

## ABSTRACT

*As technology and education intersect, the potential of Artificial Intelligence (AI) to enhance learning and engagement via chatbots is notable and has been widely acclaimed, though with mixed reports. This study thus investigates AI chatbots' effectiveness in enhancing learning outcome and engagement, using a quasi-experimental design with three null hypotheses. Participants included 90 and 96 higher education students in experimental and control groups, respectively. Data were collected using Learning Outcome Tests and Student Engagement Scales. Analysis revealed no significant difference in learning outcomes or engagement between groups, suggesting a need for a balanced integration of AI and traditional teaching methods for optimal results.*

**Keywords:** *age, artificial intelligence, chatbots, engagement, gender, learning outcome*

## INTRODUCTION

As we stand at the intersection of technological advancement and educational progress, the potential of Artificial Intelligence (AI) to improve the learning process is full of promise and intricacy. AI, known for its capability to mimic human intelligence in machines, has been distinctly utilized in education with the creation of chatbots. Chatbots are also referred to as artificial conversation entities, interactive agents, smart bots, and digital assistants (Adamopoulou & Moussiades, 2020). Chatbots are automated software applications that employ Natural Language Processing (NLP) and machine learning algorithms to interact with users through conversations. As Abu Shawar and Atwell (2015) state, it is a conversational agent that engages with users through natural language. AI chatbots provide a hopeful opportunity for personalized assistance and active involvement throughout the learning process due to their unique characteristics. These smart chatbots use natural language

processing and machine learning to interact with students in personalized ways, revolutionizing how educational content is delivered.

Chatbots have existed for a long time. In the literature of chatbots, the earliest dating is the first chatbot named ELIZA constructed in 1966. ELIZA was able to make certain kinds of natural language conversations between man and computer possible (Weizenbaum, 1966). Though it was limited in some ways, ELIZA was able to inspire the subsequent development of other chatbots (Klopfenstein et al., 2017). Other subsequent prominent examples include PARRY, ALICE, and SmarterChild (Colby et al., 1971; Essel et al., 2022; Kuhail et al., 2023). Other developments include chatbots like Siri, Alexa, Duolingo bots, Pounce, Aritz, Quizlet, Ugabot, Ada, Habitica, and Jill Watson (Adamopoulou & Moussiades, 2020; Černý, 2022; Cunningham-Nelson et al., 2019; Krohn, 2019; Labadze et al., 2023; Rathore, 2022). Recent chatbots include the likes of Siri, Alexa,

ChatGPT, Microsoft Bing, and Google Gemini, to mention a few. Chatbots can be classified based on knowledge domain, services provided, response generation method, human aid, permission, and communication channel services (Adamopoulou, & Moussiades, 2020). In the educational setting, AI chatbots can act as virtual assistants; provide students with interactive and personalized support; answer queries; facilitate learning through dynamic and responsive interactions; deliver information about classes, grades, and campus services; respond to frequently requested inquiries; assist students with their academic objectives, language acquisition, and career counseling; and speed up access to help and streamline administrative procedures (Adamopoulou & Moussiades, 2020; Hiremath et al., 2018; Mikic-Fonte et al., 2018; Min et al., 2020; Okonkwo and Ade-Ibijola, 2020; Pérez et al., 2020; Pham et al., 2018; Rathore, 2022; Sinha et al., 2020). All of these AI chatbot capabilities create a significant pathway that leads to increased engagement, retention, motivation, satisfaction, interaction, and overall general academic success for students. (Abdellatif et al., 2020; Benotti et al., 2014; Bii, 2013; Lipko, 2016; Mendoza et al., 2020; Vanichvasin, 2022).

According to scholars, since chatbots interact in a way that resembles human conversations, they are a natural fit for teaching and learning situations in which interactions between students and teachers or peers are frequent; also due to the increasing availability of chatbots, opportunities to use them in education are also increasing (Kang et al., 2022; Kim et al., 2020; Yang et al., 2019). This has led to different chatbot development for specific teaching, learning, and general educational activities by various researchers and scholars, including Clarizia et al. (2018), Neto and Fernandes (2019), Gupta et al. (2019), Kowsher et al. (2019), Winkler et al. (2020), Arruda et al. (2019), Yeves-Martínez and Pérez-Marín (2019), Tegos et al. (2019), Vijayakumar et al. (2019), Gaglo et al. (2021), Han and Lee (2022), and Ranoliya et al. (2017). Similarly, there has been a rise of chatbots that are generic in their application that have been applied to the teaching and learning process in education, both at higher and lower levels. Examples of these include ChatGPT, Microsoft Bing, and Google Gemini, to mention a few.

AI chatbots in education go beyond just sharing information; for example, they can adjust

to different learning styles by using interactive simulations for kinesthetic learners, providing detailed written explanations for reading/writing learners, and incorporating videos for visual learners. Also, they offer prompt feedback, and create a more engaging learning atmosphere. AI chatbots can enhance comprehension, retention, and application of knowledge by customizing content delivery to suit each student's pace and preferences, ultimately impacting their learning outcome. Students' learning outcome refers to the observable and measurable results of the learning process. It encompasses the knowledge, skills, and abilities that students acquire as a result of their educational experiences.

Research on the effectiveness of AI chatbots on students' performance has revealed mixed results. While some research has shown that AI interventions and tools have an effect on academic performance and learning outcome (Afonughe et al., 2021; Barbalios et al., 2013; Chin et al., 2010; Civelek et al., 2014; Fidan & Tuncel, 2019; García-Martínez et al., 2023; Ibáñez et al., 2014; Klímová & Ibna, 2023; Pareto, 2014; Pellas & Vosinakis, 2018; Shegog et al., 2012; Stieff, 2011; Tatli & Ayas, 2013; Walker et al., 2014; Winkler & Söllner, 2018; Yin et al., 2020), other studies have reported no effect of AI chatbots on academic performances and learning outcomes (Han et al., 2022; Topal, 2021; Yin et al., 2020). Although, specifically, research has shown that AI chatbots have positive effects on academic performance and learning outcome (El Shazly, 2021; Kim et al., 2021; Nghi et al., 2019).

Alongside learning outcome, students' engagement is one component that chatbots may have an influence on. Engagement in education pertains to the level of involvement, interest, and interaction that students exhibit during the learning process. It goes beyond mere participation and includes elements such as enthusiasm, focus, and sustained interest in educational activities. In the context of AI chatbots, engagement is a key parameter in evaluating how effectively these technologies captivate students' attention and foster a more interactive and immersive learning experience. Through natural and intuitive interactions, these intelligent systems may have the ability to captivate students' attention and sustain their interest throughout the learning process. The immediacy of

responses, coupled with the adaptability to individual progress, is very likely to contribute to a more responsive and engaging educational experience. Research has generally shown that AI interventions and tools have an effect on engagement and learning motivation. (Ali et al., 2023; Chen et al., 2010; Chiu et al., 2023; Hanum et al., 2023; Lee et al., 2022; Robinson & Hullinger, 2008; Sandu & Gide, 2019; Studente et al., 2020). Precisely, research has shown that AI chatbots like ChatGPT have an effect on engagement and learning motivation. (Ali et al., 2023; Hanum et al., 2023; Klímová & Ibna Seraj, 2023; Lee et al., 2022; Lin & Mubarak, 2021; Mahmoud, 2022). Although, studies by Deng and Yu (2023), and Fryer et al. (2017), report no significant effect of AI chatbots on students' engagement and interest.

Furthermore, theoretical underpinnings from the Cognitive Theory of Multimedia Learning by Richard Mayer (2009), Social Cognitive Theory by Albert Bandura (1989), and Technology Acceptance Model (TAM) by Fred Davis (1986) guided the research. The cognitive theory of multimedia learning posits that learning is most effective when information is presented in a multimedia format that optimizes the cognitive processes of selection, organization, and integration of information. In the context of this study, which was to explore the effect of AI chatbots like ChatGPT and Gemini on students' learning outcome and engagement, this theory suggests that the use of interactive multimedia platforms such as ChatGPT and Gemini can enhance learning by providing students with multiple modalities for processing information. These chatbots can deliver content through text, images, videos, and interactive simulations, catering to different learning styles and preferences. Thus, it is believed that the multimedia format engages students' attention and facilitates the construction of mental models, leading to deeper understanding and retention of course material.

Moreso, the Social Cognitive Theory also informed this study. This theory emphasizes the role of social interaction and observational learning in shaping behavior and cognition. In the context of AI chatbots, students may engage in social interactions with virtual agents, asking questions, seeking clarification, and receiving feedback. It is believed that these interactions simulate real-world

communication scenarios and are likely to promote active participation and collaborative learning.

Furthermore, the Technology Acceptance Model provided a framework for understanding students' acceptance and usage of AI chatbots in educational settings as well. According to TAM, the perceived usefulness and ease of use are key determinants of individuals' intention to adopt and utilize technology. In the context of this study, factors such as perceived usefulness in clarifying concepts, providing instant feedback, and enhancing learning outcomes, as well as perceived ease of use in interacting with the chatbots, can influence students' engagement and acceptance of these tools, thereby positively affecting learning outcomes. These theories provided further grounds that suggest that AI chatbots are likely to have an impact on students' learning outcome and engagement.

Previous research in this area shows that the potential benefits of incorporating AI chatbots into education are widely acknowledged, however, a critical gap exists in the understanding of their practical effectiveness in driving student academic achievement and sustaining engagement levels, specifically for students in higher institutions in West Africa—particularly Rivers State, Nigeria. Previous literature provides mixed results concerning the effect of AI chatbots on learning outcome and engagement from research conducted across the globe. However, there is no study, be it mixed or not, from Rivers State, Nigeria, that presents the effectiveness of AI chatbots on students' learning outcome and engagement to the best of the researchers' knowledge. It was against this background that this research was carried out to close this gap.

The following null hypotheses were tested at 0.05 alpha level:

**H0 1:** AI chatbots do not have any significant effect on students' learning outcome.

**H0 2:** AI chatbots do not have any significant effect on students engagement.

**H0 3:** The effectiveness of AI chatbots on students' learning outcome and engagement does not significantly differ between male and female students of different ages.

## METHODOLOGY

### *Research Design*

The research design was a quasi-experimental research design. Specifically, the static-group comparison design was adopted. According to Kpolovie (2014), this design is a nonequivalent posttest-only design that compares performance of an intact experimental group with that of a nonequivalent static (intact) group on the dependent variable. This study employed a static-group comparison research design, a variant of the nonequivalent posttest-only design. This research design is appropriate, as it provides opportunity to investigate the effect of the independent variable (*artificial intelligence chatbots use*) on the dependent variables (*learning outcome and engagement*) of the study. The design aimed to compare the performance of two distinct intact groups: (1) a treatment group exposed to instructional sessions using AI chatbots, and (2) a control group receiving conventional instruction without AI chatbots, using a single blind procedure as well. The comparison centered on assessing their performances, which are their learning outcome and their engagement in the Computer in Education course.

### *Participants*

The population of the study was 900 undergraduate students in year four in the 2021–2022 session offering a Computer in Education course in the Faculty of Education at the University of Port Harcourt, Rivers State, Nigeria. A multi-stage sampling procedure was employed in drawing the 184 samples and participants of the study. At the first stage, the faculty was stratified based on their department of study. Thereafter, a simple, random technique via balloting without replacement was employed to draw three departments from the seven departments in the faculty. During the second stage, a simple, random technique via balloting was employed to randomly assign the departments into experimental and control groups. Thus, intact classes constituted the experimental and control groups. Chatbots, powered by platforms like Gemini and ChatGPT, were utilized as interactive tools to facilitate learning in various ways.

### *Procedure for Data Collection*

The experimental group underwent the instructional intervention, which involved teaching the Computer in Education course utilizing AI chatbots,

specifically powered by platforms like Gemini and ChatGPT. During lectures and discussions, chatbots Gemini and ChatGPT served as virtual teaching assistants that provided additional explanations tailored to the students regarding the two content areas covered in the Computer in Education course. They also provided instant clarification on complex topics raised by students in real time during the teaching and learning class sessions in which the experiment took place. Conversely, the control group did not receive the AI chatbot-based instructional method and instead followed a standard teaching approach without AI chatbots. A single blind procedure was employed. The participants were unaware of the group to which they were assigned; however, the instructors and researchers conducting the sessions were aware of the group assignments. This blinding was crucial in preventing the participants' expectations from influencing their engagement and performance.

Ethical considerations were taken during the experimentation. Participants were thoroughly informed about the research aim and the instruments they would engage with. Participants' consent was also obtained. Privacy assurances were provided, emphasizing that all information shared would be kept confidential and would solely be utilized for research purposes. Additionally, data was coded and stored on secure computer systems, accessible exclusively to authorized researchers and safeguarded with a security protocol to prevent unauthorized access to the gathered data.

Instructional sessions for both groups were conducted separately to maximize control. Also, four sessions were held under uniform settings, ensuring similar durations and instructional materials to minimize external influences on the learning process. Following the instructional sessions, both groups underwent a posttest assessment.

The Learning Outcome Test was administered immediately following the final instructional session. Participants also completed the Engagement Scale immediately after the Learning Outcome Test. Data from both instruments were collected and anonymized through codes for analysis. The posttest assessment was designed to evaluate the participants' understanding and performance as well as their level of engagement related to the Computer in Education course material. These two tests served as the primary instruments to

measure the dependent variables: learning outcome and engagement.

#### *Instruments for Data Collection*

Two instruments were employed in obtaining the data, the Learning Outcome Test and the Engagement Scale. A single package with both instruments was divided into two parts, parts A and B. Part A elicited responses regarding demographic detail of the department of study, gender, and age; while part B was divided into two sections, A and B.

The section A Learning Outcome Test was specifically designed to measure participants' knowledge, comprehension, understanding, and proficiency related to the Computer in Education course material. The items were structured using the multiple-choice objective test format with options lettered A–D, which grade one right response with a score of 1, and three incorrect responses with a score of 0. It contained 10 items drawn from two content areas in the Computer in Education course, which are *overview of computer* and *literacy skills*. Scores obtained from the test were an indication of the outcome of their achievement in the course, as measured by their overall test score. Higher scores indicate higher outcomes, and lower scores otherwise. That means if the total score obtained in the learning outcome test, which measures knowledge and understanding of the content areas in the course (Computer in Education), then the participant has a higher learning outcome, and vice versa. The Learning Outcome Test was specifically tailored to the Computer in Education course, focusing on content areas unique to this course.

The section B Engagement Scale contained items that were designed to measure engagement level of students towards the Computer in Education course. The Engagement Scale consisted of three components: (1) *cognitive engagement*, which measures mental involvement, attention, and understanding; (2) *affective engagement*, which gauges emotional responses, interest, and enjoyment; and (3) *behavioral engagement*, which assesses active participation and interaction all towards the course. Respondents are asked to rate items on the scale using the response option of *a bit engaging*, *moderately engaging*, *very much engaging*, and *extremely engaging*, which are assigned 1, 2, 3, and 4. High scores on this instrument indicate

a high level of engagement and low scores indicate a low level of engagement. The Engagement Scale is comprised of general items related to cognitive, emotional, and behavioral engagement, and can thus be broadly applicable across different content areas. Its items were designed to capture overall student engagement, making it adaptable for various educational contexts

#### *Validity and Reliability of the Instruments*

Face and content validity were ensured by seeking expert opinions (test and measurement experts, and instructional technologist experts). These experts—who are lecturers in the field of measurement and evaluation, and also in instructional technology—critically analyzed the items for the language, content, usefulness, thoroughness, clarity, and literacy requirements. On the basis of criticisms and comments offered by experts, only those items which received significant approval of the experts were retained for item analysis. Cronbach's alpha method of internal consistency was used to estimate the instruments' reliability and construct validity. On a sample of 20 respondents, these instruments were pilot tested. Each item on the test was evaluated for quality and selection in accordance with Cronbach's alpha method. Based on the inter-item analysis and item total statistics, items were included in the final instrument. Thus, Cronbach's alpha was used to obtain a reliability coefficient of .903 for the Engagement Scale. The Learning Outcome Test reliability was established using a split-half approach to obtain a Spearman-Brown coefficient of .720.

#### *Method of Data Analysis*

Performance data obtained from the assessments of both groups were collected and analyzed statistically. Descriptive statistics of mean and Standard Deviation (SD) was employed. Thus, mean and standard deviation scores for learning outcomes and engagement levels were calculated for both the experimental and control groups. These descriptive statistics provided a basic understanding of the central tendencies and variability within the groups. A comparative analysis using appropriate statistical tests of independent samples t-tests was conducted to ascertain potential differences in the performance levels between the experimental and control groups. Data was also analyzed using a two-way Multivariate Analysis

of Variance (MANOVA) for age and gender on the learning outcome and engagement levels of the experimental group. The MANOVA allowed for the assessment of potential interactions between these two variables (age and gender) and their combined effect on the two dependent variables (learning outcome and engagement). This analysis provided insights into whether different age groups and genders responded differently to the instructional interventions. These tests were done at a chosen 0.05 alpha level.

The Statistical Package for Social Science was the software used for data analysis.

## RESULTS

**Hypothesis 1:** AI chatbots do not have any significant effect on students' learning outcome.

Table 1.

Mean, SD, and Independent Samples t-Test Analysis of Influence of AI Chatbots: No Significant Effect on Students' Learning Outcome

Groups	N	$\bar{x}$	SD	Df	T	Sig.	P	Decision
Expe.	90	7.566	1.662	184	1.914	.057	0.05	Accept $H_0$ , P > 0.05
Control	96	8.000	1.421					

Table 1 presents the analysis of the effect of AI chatbots on students' learning outcome in computer science. The table provides information on the number of participants (N) in each category, mean, standard deviation, independent samples t-test analysis, and the corresponding statistical values.

The table shows that the experimental group that was exposed to the use of AI chatbots like ChatGPT and Gemini in the process of teaching the course Computer in Education had a mean score of 7.566 and an SD of 1.662. The control group that was not exposed to the use of AI chatbots like ChatGPT and Gemini in the process of teaching the course Computer in Education had a mean score of 8.000 and an SD of 1.421. From the table, the experimental group had a lesser mean than the control group. This shows that the control group had a higher learning outcome score than the experimental group that was exposed to AI chatbots like ChatGPT and Gemini in the process of teaching the course Computer in Education

The independent samples t-test analysis is conducted to determine if there is a significant difference between the learning outcome of the students that were exposed to the use of AI chatbots (experimental group) and those that were not (control group). The t-value is calculated as 1.914, with a degree of freedom (*df*) of 184. The significance value (Sig.) is reported as .057. Based on the results of the t-test, the null hypothesis is accepted ( $H_0$ ) at the 0.05 significance level, that is  $t = 1.914$ ,  $p = .057$  ( $p > .005$ ). This indicates that there is no significant difference between the learning outcome of students who were exposed to AI chatbots and those who were not ( $184 = 1.914$ ,  $p > .05$ ). This means that the effect of AI chatbots on students' learning outcome is not significant.

**Hypothesis 2:** AI Chatbots do not have any significant effect on students' engagement.

Table 2.

Mean, SD, and Independent Samples t-Test Analysis of Influence of AI Chatbots: No Significant Effect on Students' Engagement

Groups	N	$\bar{x}$	SD	Df	T	Sig.	P	Decision
Expe.	90	50.77	9.552	184	1.271	.205	0.05	Accept $H_0$ , P > 0.05
Control	96	52.58	9.798					

Table 2 presents the analysis of the effect of AI chatbots on students' engagement in computer science. The table shows that the experimental group exposed to the use of AI chatbots like ChatGPT and Gemini in the process of teaching the course Computer in Education had an engagement mean score of 50.77 and an SD of 9.552. The control group not exposed to the use of AI chatbots like ChatGPT and Gemini in the process of teaching the course Computer in Education had an engagement mean score of 52.58 and an SD of 9.789. From the table, the experimental group had a lesser mean than the control group. This shows that the control group had a higher engagement score than the experimental group exposed to AI chatbots like ChatGPT and Gemini in the process of teaching the course Computer in Education

The independent samples t-test analysis conducted shows a t-value calculated as 1.271,

with a degree of freedom (*df*) of 184. The significance value (Sig.) is reported as .205. Based on the results of the t-test, the null hypothesis is accepted ( $H_0$ ) at the 0.05 significance level, that is  $t = -1.271$ ,  $p = .205$  ( $p > .005$ ). This indicates that there is no significant difference between the engagement of students who were exposed to AI chatbots and those who were not ( $184$ ) =  $-1.914$ ,  $p > .05$ . This means that AI chatbots effect on students' engagement is not significant.

Hypothesis 3: The effectiveness of AI chatbots on students' learning outcome and engagement does not significantly differ between male and female students of different ages.

Table 3 shows the mean representing the learning outcome and engagement scores for male

and female students of different ages exposed to AI chatbots in the learning the course Computer in Education. It shows that males had a learning outcome score of 7.000, while females had a learning outcome score of 7.796. This shows that the AI chatbots were more effective for female students who had a higher mean score than male. However, this difference in mean score statistically is not significant as  $p = .103 > 0.05$ . The table also shows that students aged 31 years and older had the highest learning outcome scores (10.000), followed by students 21–25 years old (7.727), and then 26–30 years old (7.000). Also, this difference in the means scores of the ages of the respondents was not significant as  $p = .101 > 0.05$ . The table reveals that AI chatbots had the most effect on the

Table 3.

Mean, SD, and Two-Way MANOVA Analysis of Effectiveness of AI Chatbots on Students' Learning Outcome and Engagement for Male and Female Students of Different Ages

Variables	Gender	Age	Mean	Std. Deviation	N	Univariate test(f)	Sig	Multivariate Test(f)	Sig
Learning	Male	21–25 yrs.	6.9286	1.73046	14				
		26–30 yrs.	6.8182	1.77866	11				
		31 yrs. and older	10.0000	.	1				
		Total	7.0000	1.78885	26				
	Female	21–25 yrs.	7.9423	1.63795	52				
		26–30 yrs.	7.1667	1.02986	12				
		Total	7.7969	1.56530	64	2.718	.103		
	Total	21–25 yrs.	7.7273	1.69656	66				
		26–30 yrs.	7.0000	1.41421	23				
		31 yrs. and older	10.0000	.	1				
		Total	7.5667	1.66288	90	2.360 .330	.101 .423		
	Engagement	Male	21–25 yrs.	50.6429	12.19462	14			
26–30 yrs.			53.5455	8.73343	11				
31 yrs. and older			47.0000	.	1				
Total			51.7308	10.52828	26			.413	.663
Female		21–25 yrs.	50.3846	8.92925	52				
		26–30 yrs.	50.4167	10.66394	12				
		Total	50.3906	9.18698	64	.459	.500		
Total		21–25 yrs.	50.4394	9.60790	66				
		26–30 yrs.	51.9130	9.69964	23				
		31 yrs. and older	47.0000	.	1				
		Total	50.7778	9.55267	90	.299 .648.	.742 .567		

learning outcome scores of male students aged 31 years and older (10.00), followed by female students aged 21–25 years old (7.9423), then female students aged 26–30 years old (7.166), then male students 21–25 years old (6.928), and lastly male students aged 26–30 years old (6.818). The univariate test shows that this difference is statistically not significant as  $F(2, 89) = .330$ ,  $P > .05$ , i.e.  $p = .423$ , i.e.  $p = .423$ . This implies that the difference in the learning outcome scores of male and female students aged 21–25 years, 26–30 years, and older than 30 statistically is not significant, as the univariate test shows no significant difference  $F(2, 89) = .330$ ,  $P > .05$ , i.e.  $p = .423$ , i.e.  $p = .423$ .

It can also be observed from Table 3 that male had an engagement mean score of 51.73, while female had an engagement score of 50.39. This shows that the AI chatbots were more effective for male students who had a higher mean score than female. However, this difference in mean score statistically is not significant as  $p = .500 > 0.05$ . The table also shows that students aged 26–30 years (51.91) had the highest engagement scores, followed by students 21–25 years (50.43), and then 30 years or older (47.00). Also, this difference in the means scores of the ages of the respondents was not significant according to the univariate test as  $F(2, 89) = .299$ ,  $P > .05$ , i.e.  $p = .742$ , i.e.  $p = .742$ . The table reveals that the AI chatbots had the most effect on the engagement of male students aged 26–30 years (53.54), followed by male students aged 21–25 years (50.64), then by female students aged 26–30 years (50.54), female students aged 21–25 years (50.38), and then males 30 years or older (47.00). The univariate test shows that this difference is statistically not significant as  $F(2, 89) = .642$ ,  $P > .05$ , i.e.  $p = .527$ , i.e.  $p = .527$ . This implies that the difference in the engagement scores of male and female students aged 21–25 years, 26–30 years, and 30 years or older statistically is not significant, as the univariate test shows no significant difference  $F(2, 89) = .642$ ,  $P > .05$ , i.e.  $p = .527$ , i.e.  $p = .527$ .

The two-way MANOVA multivariate test examines the overall effectiveness of AI chatbots on students' learning outcome and engagement for male and female students of different ages. The results show no significant difference in the influence of age and gender on the learning outcome and engagement of students exposed to AI chatbots ( $f = .416$ ,  $p = .663$ ). That is the learning outcome and

engagement level of male and female students aged 21–25 years, 26–30 years, and 30 years and older does not differ significantly as seen by the Wilk's  $\Lambda = .416$ ,  $p = .663$ ,  $> .05$ . This simply connotes that the age (whether one is between age 21–25, 26–30, or 30 years or older) and gender (whether one is male or female) does not have a significant influence on how effective AI chatbots will be on learning outcome and engagement. Therefore, the hypothesis of effectiveness of AI chatbots on students' learning outcome and engagement being not significantly different for male and female students of different ages is accepted, and the alternate rejected.

## DISCUSSION OF FINDINGS

This indicates that there is no effect of AI chatbots on students' learning outcome. This is in tandem with the study of Yin et al. (2020), Topal et al. (2021), and Han et al. (2022), who did not find any significant difference between the experimental and control group based on the effectiveness of chatbots on the learning performance and outcome of students. It is, however, not in agreement with that of El Shazly (2021), Kim et al. (2021), and Nghi et al. (2019). The reason for the divergent findings could be due to the difference in the specific types of AI chatbots and quasi-experimental designs that were adopted; the difference in the measurement instruments used in obtaining data; and also the difference in the course that learning outcomes were obtained from. For instance, difference between the current study's result and that of El-Shalzy (2021) could be due to the fact that the latter focused on speaking performance and applied a different type of quasi-experimental research design, and also used role-play as a method of assessment, while the current study focused on learning outcome, used a learning outcome test in Computer in Education that was designed using the multiple-choice objective test format.

The results show that the control group has a higher learning outcome score than the experimental group. The control group having a higher learning outcome score in comparison to the experimental group can be due to factors like distractions and overreliance on technology. While the intention behind integrating chatbots was to enhance learning, students in the experimental group might have become overly reliant on the

chatbots for information retrieval and clarification. This overdependence on technology could lead to distractions during class discussions or reduced engagement with course material beyond what the chatbots provided. In contrast, the control group, not relying on chatbots, might have engaged more actively with the content, leading to a deeper understanding and higher learning outcomes. Also, the quality of the interaction could be another factor.

Despite the convenience of instant responses from chatbots, the quality of interaction might not have been as effective as anticipated. Chatbots, while capable of providing information, may lack the understanding and personalized feedback that human instructors can offer. Students in the control group, receiving direct guidance and explanations from instructors, could have benefited from more tailored support and deeper insights into the course material, resulting in higher learning outcomes. Technical issues is another factor to consider. Technical issues with the chatbot platforms—such as glitches, difficulty in accessing responses, and fluctuating or poor network connectivity—could have disrupted the learning process for the experimental group. In contrast, the control group, not facing these technological hurdles, might have experienced smoother and more focused learning experiences, contributing to their higher learning outcome scores.

The absence of a significant impact of AI chatbots like ChatGPT and Gemini on the learning outcome test representing the achievement scores of students in the Computer in Education course may stem from various practical factors. The pedagogical approach in the control group might have been robust enough to yield comparable outcomes, mitigating the potential advantages offered by the ChatGPT and Gemini AI chatbots. Additionally, the lack of statistical significance may also be influenced by external variables, such as lack of adaptability and personalization in the chatbots' interactions, which could have limited their ability to cater to the diverse learning styles of students, resulting in a generic and less impactful educational experience. Also, the quality of content delivered by the AI chatbots, along with the instructional methods employed, may not have sufficiently engaged students or facilitated a deeper understanding of the course material. This was emphasized by Vanichvasin (2022) who

asserts in his study that findings suggest that chatbot technology seem to have more advantages than disadvantages in supporting teaching and learning if it is properly designed and developed with educational intentionality. Also, inadequate training for both the instructors and students on how to maximize the potential of these chatbots could have hindered their effective use; the study's time-frame might have been too short or lacked the necessary intensity for the chatbots to demonstrate a measurable influence on academic performance; and individual differences among students might have also contributed. All these could have masked the true impact of the AI chatbots on student's learning outcome.

This results suggest that the integration of AI chatbots in the teaching of computer in education may not be providing a discernible advantage over traditional teaching methods. This challenges the assumption that technology-driven approaches necessarily lead to superior learning outcomes. Furthermore, the results emphasize the importance of considering the holistic educational experience, encompassing factors beyond the immediate impact of technology, such as teacher–student interactions, pedagogical methods, assessment tools, and the overall learning environment. In the absence of a significant difference, it is suggested that a thoughtful balance between technological integration and traditional teaching practices may be crucial to achieving optimal learning outcomes.

The results also indicate that there is no effect of AI chatbots on students' engagement. This is in tandem with the studies of Deng and Yu (2023), and Fryer et al. (2017), whose investigations found no significant effect of AI chatbots on students' engagement. However, it is not in agreement with that of Abbas (2022), Klímová and Ibna (2023), Lee et al. (2022), Lin and Mubarak (2021), and Mahmoud (2022), whose studies found significant effects of AI chatbots on students' engagement. The reason for the divergent finding could be due to the difference in the treatment (difference in the chatbots applied), the nature of the quasi-experimental designs that were adopted, and the possible difference in the measurement instruments used in obtaining on the engagement. Furthermore, the current studies, along with studies being reviewed and analyzed, were conducted in different global climates.

The absence of a significant effect of AI chatbots like ChatGPT and Gemini on the engagement level of students in the Computer in Education course may stem from various practical factors. The quality and effectiveness of the AI chatbot interactions may not have sufficiently captured students' attention or interest, and content delivered by the Gemini and ChatGPT chatbots on the Computer in Education course may have lacked relevance, interactivity, and personalization. Another possible reason could be the novelty effect wearing off over time. Initially, students exposed to AI chatbots might have found the technology intriguing; but as the novelty wore off, it might not have sustained a higher level of engagement compared to the control group. Moreover, the adaptability of AI chatbots to diverse learning styles may have been limited, resulting in a less personalized learning experience. If the chatbots failed to cater to individual preferences and varied learning needs, the potential impact on engagement might have been diminished. The study design, the assessment tool used to measure engagement, the duration of exposure to AI chatbots, and other factors' influences would have mitigated against any significant effect that AI chatbots would have had on students' engagement.

The implication of the results indicating no significant difference in engagement scores between the control group and the experimental group suggests that the integration of AI chatbots, in its current form, did not lead to a measurable enhancement in student engagement compared to traditional teaching methods. This challenges the assumption that technological interventions inherently result in increased engagement. The lack of statistical significance implies that, despite potential variations in mean scores, these differences are not significant enough to attribute to the influence of AI chatbots.

Furthermore, the results underscore the importance of a diverse approach to technology integration in education. The results highlight the need for continuous evaluation and refinement of AI applications to ensure they align with the dynamic nature of student engagement and learning preferences. Educators and institutions may need to critically assess the relevance and effectiveness of AI chatbots in promoting sustained and meaningful engagement over time. In practical terms, this

result encourages a reevaluation of the role of AI in enhancing student engagement, prompting educators to consider adjustments in the design, content, and delivery of AI chatbots to better align with the diverse needs and preferences of students. It also emphasizes the importance of recognizing the multifaceted nature of engagement, encompassing both technological and traditional teaching elements to create an optimal learning environment.

The results also show no significant difference in the influence of age and gender on the learning outcome and engagement of students exposed to AI chatbots. This aligns with the findings of Essel et al. (2022), which indicate no significant difference related to gender and age on the academic performance of students exposed to the chatbot. Same also with Sandu and Gide (2019), who also report no significant difference between gender and age and the adoption of a chatbot. The findings suggest that the influence of age and gender on learning outcome and engagement was not significantly different. This homogeneity across age groups and genders might indicate that AI chatbots had a consistent, non-discriminatory impact on engagement, affecting both male and female students of between 21–25 years, 26–30 years, and 30 years and older brackets. Furthermore, other external variables influence within the educational setting may have neutralized any significant distinctions that might have been tied to age or gender of the students.

## RECOMMENDATIONS

Given the results, several practical recommendations emerged. Educators should improve the content delivered by AI chatbots to ensure it is relevant, interactive, and captures students' interest, for instance, by using a multimodal presentation. Utilizing a variety of media formats, such as text, images, videos, and interactive simulations, to present content through the chatbots can make these contents more relevant, interactive, engaging, and more improved. Tailoring the content to the course objectives and incorporating engaging elements can also contribute to increased student engagement in and out of the course and learning outcome as well. Also, educators should be mindful of the potential novelty effects associated with the introduction of AI technology by recognizing that initial interest may not always translate into sustained engagement, and useful design interventions may

be needed to maintain students' curiosity and interest over time. Furthermore, educational institutions should provide training for educators, and educators also can seek training on how to effectively integrate AI and AI chatbots into their assessment and teaching methods. They should also seek collaboration between AI chatbot developers to create a synergy that optimizes the benefits of both traditional and technological approaches.

It is also important for educators to explore the possibility of customizing the use of AI chatbots based on the specific requirements of different subjects or topics. Tailoring their application to areas where they can provide the most benefit may lead to more pronounced positive effects. It is important that educators, and educational institutions implement mechanisms for collecting feedback from students regarding their experiences with AI chatbots and utilizing this feedback to make real-time adjustments and improvements, ensures that the technology remains responsive to students' engagement needs.

## **CONCLUSION**

AI chatbots like ChatGPT and Gemini applications in learning did not have a significant effect on students' learning outcome and engagement. The study was limited in the use of the static group quasi-experimental design and a short intervention time for the experimental groups. It is suggested that further studies can be done using an experimental design that allows for pretest and posttest comparison, a more individualized and custom-made chatbot, as well as a longer intervention time for the experimental group. It is also suggested that a follow-up study be conducted with the influence of such factors as respondents' computer abilities and comfort using computer programs factored into the research, which could have a significant impact on the study.

## References

- Abbas, N., Whitfield, J., Atwell, E., Bowman, H., Pickard T., & Walker A. (2022). Online chat and chatbots to enhance mature student engagement in higher education. *International Journal of Lifelong Education*, 41(3), 308–326. <https://doi.org/10.1080/02601370.2022.2066213>
- Ali, J. K. M., Shamsan, M. A. A., Hezam, T. A., & Mohammed, A. A. Q. (2023). Impact of ChatGPT on learning motivation: teachers and students' voices. *Journal of English Studies in Arabia*, 2(1), 41–49. <http://dx.doi.org/10.56540/jesaf.v2i1.51>.
- Abdellatif, A. & Costa, D., Badran, K., Abdalkareem, R. & Shihab, E. (2020). Challenges in chatbot development; A study of stack overflow posts. In Proceedings of the 17<sup>th</sup> international Conference on Mining Software Repositories (pp.174-185). <https://doi.org/10.1145/3379597.3387472>.
- Al-Abdullatif, A. M., Al-Dokhny, A. A., & Drwish, A. M. (2023). Implementing the Bashayer chatbot in Saudi higher education: Measuring the influence on students' motivation and learning strategies. *Frontier Psychology*, 14, 1129070. <https://doi.org/10.3389/fpsyg.2023.1129070>
- Afonughe, E., Onah, E. N., Uzoma, A. C., Andor, S. E., & Orisakwe, C. U. (2021). Integration of artificial intelligence tool (AI chatbot) into teaching and learning: A panacea for improving universities' educational and administrative duties in South-South, Nigeria. *Journal of Computer Science and Systems Biology*, 14(6), 357.
- Abu Shawar, B., & Atwell, E. (2007). Chatbots: Are they really useful? *Journal for Language Technology and Computational Linguistics*, 22(1), 29–49. <http://dx.doi.org/10.21248/jlcl.22.2007.88>
- 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>
- Arruda, D., Marinho, M., Souza, E., & Wanderley, F. (2019). A chatbot for goal-oriented requirements modeling. In Misra, S. (Ed.), *Computational Science and Its Applications—ICCSA 2019. Lecture Notes in Computer Science* (pp. 506–519). Springer. [http://dx.doi.org/10.1007/978-3-030-24305-0\\_38](http://dx.doi.org/10.1007/978-3-030-24305-0_38)
- Barbalios, N., Ioannidou, I., Tzionas, P., & Paraskeuopoulos, S. (2013). A model supported interactive virtual environment for natural resource sharing in environmental education. *Computers & Education*, 62, 231–248. <https://doi.org/10.1016/j.compedu.2012.10.029>
- Bandura, A., & National Institute of Mental Health. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall, Inc.
- Benotti, L., Martínez, M. C., & Schapachnik, F. (2014). Engaging high school students using chatbots. In Proceedings of the 2014 Innovation and Technology in Computer Science Education Conference (pp. 63–68). ACM, New York, USA. <https://doi.org/10.1145/2591708.2591728>
- Bii, P. K. (2013). Chatbot technology: A possible means of unlocking student potential to learn how to learn. *Educational Research*, 4(2), 218–221.
- Chen, P. D., Lambert, A. D., & Guidry, K. R. (2010). Engaging online learners: The impact of web-based learning technology on college student engagement. *Computers & Education*, 54(4), 1222–1232. <https://doi.org/10.1016/j.compedu.2009.11.008>
- Chiu, T. K. F., 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. <http://dx.doi.org/10.1080/10494820.2023.2172044>
- Chin, D. B., Dohmen, I. M., Cheng, B. H., Oppezzo, M. A., Chase, C. C., & Schwartz, D. L. (2010). Preparing students for future learning with teachable agents. *Educational Technology Research and Development*, 58(6), 649–669. <https://doi.org/10.1007/s11423-010-9154-5>
- Clarizia, F., Colace, F., LomGeminii, M., Pascale, F., & Santaniello, D. (2018). Chatbot: An education support system for student. In *International Symposium on Cyberspace Safety and Security*, (pp. 291–302). [http://dx.doi.org/10.1007/978-3-030-01689-0\\_23](http://dx.doi.org/10.1007/978-3-030-01689-0_23)
- Civelek, T., Ucar, E., Ustunel, H., & Aydin, M. K. (2014). Effects of a haptic augmented simulation on K–12 students' achievement and their attitudes towards physics. *Science and Technology Education*, 10(6), 565–574. <https://doi.org/10.12973/eurasia.2014.1122a>
- Černý, M. (2022). The history of chatbots: The journey from psychological experiment to educational object. *Journal of Applied Technical and Educational Sciences*, 12(3), 322–322. <https://doi.org/10.24368/jates322>
- Colby, K. M., Weber, S., & Hilf, F. D. (1971). Artificial paranoia. *Artificial Intelligence*, 2(1), 1–25. [http://dx.doi.org/10.1016/0004-3702\(71\)90002-6](http://dx.doi.org/10.1016/0004-3702(71)90002-6)
- Cunningham-Nelson, S., Boles, W., Trouton, L., & Margerison, E. (2019). A review of chatbots in education: Practical steps forward. In 30th Annual Conference for the Australasian Association for Engineering Education (AAEE 2019): *Educators Becoming Agents of Change: Innovate, Integrate, Motivate*, (pp. 299–306). Engineers Australia. Retrieved from [https://aaee.net.au/wp-content/uploads/2020/07/AAEE2019\\_Annual\\_Conference\\_paper\\_184.pdf](https://aaee.net.au/wp-content/uploads/2020/07/AAEE2019_Annual_Conference_paper_184.pdf)
- Deng, X., & Yu, Z. A. (2023). Meta-analysis and systematic review of the effect of chatbot technology use in sustainable

- education. *Sustainability*, 15(4), 2940. <https://doi.org/10.3390/su15042940>
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
- 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). <https://doi.org/10.1186/s41239-022-00362-6>
- El Shazly, R. (2021). Effects of artificial intelligence on English speaking anxiety and speaking performance: A case study. *Expert Systems*, 38(3), 1–15. <https://doi.org/10.1111/exsy.12667>
- Fidan, M., & Tuncel, M. (2019). Integrating augmented reality into problem based learning: The effects on learning achievement and attitude in physics education. *Computers & Education*, 142, 103635. <https://doi.org/10.1016/j.compedu.2019.103635>
- Fryer, L. K., Ainley, M., Thompson, A., Gibson, A., & Sherlock, Z. (2017). Stimulating and sustaining interest in a language course: An experimental comparison of chatbot and human task partners. *Computers in Human Behavior*, 75, 461–468. <https://doi.org/10.1016/j.chb.2017.05.045>
- Gaglo, K., Degboe, B. M., Kossingou, G. M., & Ouya, S. (2021). Proposal of conversational chatbots for educational remediation in the context of Covid-19. In 2021 23rd International Conference on Advanced Communication Technology (ICACT), 23, (pp. 354–358). <http://dx.doi.org/10.23919/ICACT51234.2021.9370946>
- García-Martínez, I., Fernández-Batanero, J. M., Fernández-Cerero, J., & León, S. P. (2023). Analysing the impact of artificial intelligence and computational sciences on student performance: Systematic review and meta-analysis. *Journal of New Approaches in Educational Research*, 12(1), 171–197. <https://doi.org/10.7821/naer.2023.1.1240>
- Gupta, S., Jagannath, K., Aggarwal, N., Sridar, R., Wilde, S., & Chen, Y. (2019). Artificially Intelligently (AI) tutors in the classroom: A need assessment study of designing chatbots to support student learning. Twenty-Third Pacific Asia Conference on Information Systems, July 8–12, 2019, Xi'an, China.
- Han, J.-W., Park, J., & Lee, H. (2022). Analysis of the effect of an artificial intelligence chatbot educational program on non-face-to-face classes: A quasi-experimental study. *BMC Medical Education*, 22, 830. <https://doi.org/10.1186/s12909-022-03898-3>
- Heffernan, C. J. (1988). Social foundations of thought and action: A social cognitive theory. Albert Bandura Englewood Cliffs, New Jersey: Prentice Hall. *Behaviour Change*, 5(1), 37–38. <https://doi.org/10.1017/S0813483900008238>
- Han, S., & Lee, M. K. (2022). FAQ chatbot and inclusive learning in massive open online courses. *Computers and Education*, 179, 104395. <https://doi.org/10.1016/j.compedu.2021.104395>
- Hanum Siregar, F., Hasmayni, B., & Lubis, A. H. (2023). The analysis of ChatGPT usage impact on learning motivation among scout students. *International Journal of Research and Review*, 10(7), 632–638. <https://doi.org/10.52403/ijrr.20230774>
- Hiremath, G., Hajare, A., Bhosale, P., Nanaware, R., & Wagh, K. S. (2018). Chatbot for education system. *International Journal of Advance Research, Ideas and Innovations in Technology*, 4(3), 37–43.
- Huang, A., Lu, O., & Yang, S. J. H. (2023). Effects of artificial intelligence—Enabled personalized recommendations on learners' learning engagement, motivation, and outcomes in a flipped classroom. *Computers & Education*, 194(6), 104684. <http://dx.doi.org/10.1016/j.compedu.2022.104684>
- Ibáñez, M. B., Serio, A. D., Villarán, D., & Kloos, C. D. (2014). Experimenting with electromagnetism using augmented reality: Impact on flow student experience and educational effectiveness. *Computers & Education*, 71, 1–13. <https://doi.org/10.1016/j.compedu.2013.09.004>
- Kang, K.-A., Kim, S.-J., & Kang, S. R. (2022). Elementary school students' awareness of the use of artificial intelligence chatbots in violence prevention education in South Korea: A descriptive study. *Child Health Nursing Research*, 28(4), 291–298. <https://doi.org/10.4094/chnr.2022.28.4.291>
- Klopfenstein, L., Delpriori, S., Malatini, S., & Bogliolo, A. (2017). The rise of bots: A survey of conversational interfaces, patterns, and paradigms. In Proceedings of the 2017 Conference on Designing Interactive Systems (pp. 555–565). Association for Computing Machinery. <http://dx.doi.org/10.1145/3064663.3064672>
- Kim, H. J., Yoo, S., & Kim, E.-y. (2020). AI chatbot design and development model for classroom instruction: Designing learning conversations. *Brain, Digital, & Learning*, 10(3), 535–545. <https://doi.org/10.31216/BDL.2020.10.3.535>
- Klímová, B., & Ibna Seraj, P. M. (2023). The use of chatbots in university EFL settings: Research trends and pedagogical implications. *Frontiers in Psychology*, 14, 1131506. <http://dx.doi.org/10.3389/fpsyg.2023.1131506>
- Kpolovie, P. J. (2014). *Advance Research Methods*. New Owerri: Springfield Publishers Ltd.
- Krohn, F. (2019). Chatbot learning: Real-world applications and educational possibilities. *Journal of Technology and Teacher Education*, 27(4), 437–457.
- Kowsher, M., Tithi, F. S., Alam, M. A., Huda, M. N., Moheuddin, M. M., & Rosul, M. G. (2019). Doly: Bengali chatbot for Bengali

- education. In 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), (pp. 1–6). <http://dx.doi.org/10.1109/ICASERT.2019.8934592>
- Kim, H. S., Cha, Y., & Kim, N. Y. (2021). Effects of AI chatbots on EFL students' communication skills. *Korean Journal of English Language and Linguistics*, 21, 712–734. <https://doi.org/10.15738/kjell.21.202108.712>
- Kuhail, M. A., Alturki, N., Alramlawi, S., & Alhejori K. (2023). Interacting with educational chatbots: A systematic review. *Education and Information Technologies*, 28, 973–1018. <https://doi.org/10.1007/s10639-022-11177-3>
- Lin, C. J., & Mubarak, H. (2021). International forum of educational technology & society learning analytics for investigating the mind map-guided AI chatbot approach in an EFL flipped speaking classroom. *Technology in Society*, 24, 16–35.
- Labadze, L., Grigolia, M., & Machaidze, L. (2023). Role of AI chatbots in education: Systematic literature review. *International Journal of Educational Technology in Higher Education*, 20(1), 1–17. <https://doi.org/10.1186/s41239-023-00426-1>
- Lipko, H. (2016, November 10). Meet Jill Watson: Georgia Tech's first AI teaching assistant. Georgia Institute of Technology website. Retrieved from <https://pe.gatech.edu/blog/meet-jill-watson-georgia-techs-first-ai-teaching-assistant>
- Mayer, R. E. (2009). *Multimedia Learning* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511811678>
- Mendoza, S., Hernández-León, M., Sánchez-Adame, L. M., Rodríguez, J., Decouchant, D., & Meneses Viveros, A. (2020). Supporting student-teacher interaction through a chatbot. In *Learning and Collaboration Technologies, Human and Technology Ecosystems*, (pp. 93–107). [http://dx.doi.org/10.1007/978-3-030-50506-6\\_8](http://dx.doi.org/10.1007/978-3-030-50506-6_8)
- Mahmoud, R. H. (2022). Implementing AI-based conversational chatbots in EFL speaking classes: An evolutionary perspective. *Research Square*, July 2022. <https://doi.org/0.21203/rs.3.rs-1911791/v1>
- Min, Y., Ahn, J., & Kim, S. (2020). Structural relationship analysis between the intention to use educational chatbots and influential factors based on technology acceptance model. *Journal of Korean Association for Educational Information and Media*, 26(4), 799–825. <https://doi.org/10.15833/KAFEIAM.26.4.799>
- Mikic-Fonte, F. A., Llamas-Nistal, M., & Caeiro-Rodríguez, M. (2018). Using a chatterbot as a FAQ assistant in a course about computers architecture. In 2018 IEEE Frontiers in Education Conference (FIE), October 2018, (pp. 1–4).
- Natale, S. (2019). If software is narrative: Joseph Weizenbaum, artificial intelligence and the biographies of ELIZA. *New Media and Society*, 21(3), 712–728. <https://doi.org/10.1177/1461444818804980>
- Nghi, T. T., Phuc, T. H., & Thang, N. T. (2019). Applying AI chatbot for teaching a foreign language: Empirical research. *International Journal of Scientific & Technology Research*, 8(11), 897–902.
- Neo, M., Lee, C. P., Tan, H. Y.-J., Neo, T. K., Tan, Y. X., Mahendru, N., & Ismat, Z. (2022). Enhancing students' online learning experiences with Artificial Intelligence (AI): The MERLIN Project. *International Journal of Technology*, 13(5), 1023–1034. <https://doi.org/10.14716/ijtech.v13i5.5843>
- Neto, A. J. M., & Fernandes, M. A. (2019). Chatbot and conversational analysis to promote collaborative learning in distance education. In 2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT), 21(61), (pp. 324–326). <http://dx.doi.org/10.1109/ICALT.2019.00102>
- Okonkwo, C. W., & Ade-Ibijola, A. (2020). Python-Bot: A chatbot for teaching python programming. *Engineering Letters*, 29(1), 25–34.
- Okonkwo, C. W. & Ade-Ibijola, A. (2021). Chatbots applications in education: A systematic review. *Computers and Education Artificial Intelligence*, 2(2), 100033. <http://dx.doi.org/10.1016/j.caeai.2021.100033>
- Lee, Y.-F., Hwang, G.-J., & Chen, P.-Y. (2022). Impacts of an AI-based chatbot on college students' after-class review, academic performance, self-efficacy, learning attitude, and motivation. *Educational Technology Research and Development*, 70(5), 1843–1865. <http://dx.doi.org/10.1007/s11423-022-10142-8>
- Pellas, N., & Vosinakis, S. (2018). The effect of simulation games on learning computer programming: A comparative study on high school students' learning performance by assessing computational problem-solving strategies. *Education and Information Technologies*, 23(3), 2423–2452. <https://doi.org/10.1007/s10639-018-9724-4>
- Pérez, 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>
- Pareto, L. (2014). A teachable agent game engaging primary school children to learn arithmetic concepts and reasoning. *International Journal of Artificial Intelligence in Education*, 24(3), 251–283. <https://doi.org/10.1007/s40593-014-0018-8>
- Pham, X.P., Pham, T. Nguyen, Q.M, Nguyen, T. N. & Huong, C. (2018). Chatbot as an Intelligent Personal Assistant for Mobile Language Learning. In *Proceedings of the 2018 2nd International Conference on Education and E-Learning* (pp. 16–21). <https://doi.org/10.1145/3291078.3291115>
- Rathore, B. (2022). Exploring the potential impacts of chatbot software/apps (ChatGPT) on education: Benefits, drawbacks,

- and future prospects. *International Journal of Multidisciplinary Innovation and Research Methodology* 1(1), 39–47.
- Ranoliya, B. R., Raghuwanshi, N., & Singh, S. (2017). Chatbot for university related FAQs. In 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), September 2017, (pp. 1525–1530). IEEE Publications.
- Robinson, C., & Hullinger, H. (2008). New benchmarks in higher education: Student engagement in online learning. *Journal of Education for Business*, 84(2), 101–109. <https://doi.org/10.3200/JOEB.84.2.101-109>
- Sinha, S., Basak, S., Dey, Y., & Mondal, A. (2020). An educational chatbot for answering queries. In Mandal, J. K., & Bhattacharya, D. (Eds.), *Advances in Intelligent Systems and Computing*, (pp. 55–60). Springer. [http://dx.doi.org/10.1007/978-981-13-7403-6\\_7](http://dx.doi.org/10.1007/978-981-13-7403-6_7)
- Sandu, N., & Gide, E. (2019). Adoption of AI-Chatbots to enhance student learning experience in higher education in India. In 18th International Conference on Information Technology Based Higher Education and Training (ITHET), September 2019, (pp. 1–5). IEEE Publications.
- Shegog, R., Lazarus, M. M., Murray, N. G., Diamond, P. M., Sessions, N., & Zsigmond, E. (2012). Virtual transgenics: Using a molecular biology simulation to impact student academic achievement and attitudes. *Research in Science Education*, 42(5), 875–890. <https://doi.org/10.1007/s11165-011-9216-7>.
- Stieff, M. (2011). Improving representational competence using molecular simulations embedded in inquiry activities. *Journal of Research in Science Teaching*, 48(10), 1137–1158. <https://doi.org/10.1002/tea.20438>
- Studente, S., Ellis, S., & Garivaldis, S. (2020). Exploring the potential of chatbots in higher education: A preliminary study. *International Journal of Educational and Pedagogical Sciences*, 14(9), 768–771.
- Tatli, Z., & Ayas, A. (2013). Effect of a virtual chemistry laboratory on students' achievement. *Educational Technology & Society*, 16(1), 159–170.
- Topal, A. D., Eren, C. D., & Geçer, A. K. (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>
- Tegos, S., Psathas, G., Tsiatsos, T., & Demetriadis, S. (2019). Designing conversational agent interventions that support collaborative chat activities in MOOCs. In EMOOCs 2019: Work in Progress Papers of the Research, Experience and Business Tracks. CEUR Workshop Proceedings, May 20–22, 2019, Naples, Italy. <http://ceur-ws.org/Vol-2356/>
- Vijayakumar, R., Bhuvaneshwari, B., Adith, S., & Deepika, M. (2019). AI based student bot for academic information system using machine learning. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 5(2), 590–596. <https://doi.org/10.32628/CSEIT1952171>
- Vanichvasin, P. (2022). Impact of chatbots on student learning and satisfaction in the entrepreneurship education programme in higher education context. *International Education Studies*, 15(6), 15. <https://doi.org/10.5539/ies.v15n6p15>
- Walker, E., Rummel, N., & Koedinger, K. R. (2014). Adaptive intelligent support to improve peer tutoring in algebra. *International Journal of Artificial Intelligence in Education*, 24(1), 33–61. <https://doi.org/10.1007/s40593-013-0001-9>
- 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., Hobert, S., Salovaara, A., Söllner, M., Leimeister, J. M. (2020). Sara, the lecturer: improving learning in online education with a scaffolding-based conversational agent. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, (pp. 1–14). <https://doi.org/10.1145/3313831.3376781>
- Winkler, R., & Söllner, M. (2018). Unleashing the potential of chatbots in education: A state-of-the-art analysis. In 2018 Academy of Management Annual Meeting (AOM), 1, 15903. <http://dx.doi.org/10.5465/AMBPP.2018.15903abstract>
- Yin, Q., & Satar, M. (2020). English as a foreign language learner interaction with chatbots: Negotiation for meaning. *International Online Journal of Education and Teaching (IOJET)*, 7(2), 390–410. <http://iojet.org/index.php/IOJET/article/view/707>
- Yeves-Martínez, P., & Pérez-Marín, D. (2019). Prof. Watson: A pedagogic conversational agent to teach programming in primary education. In *Multidisciplinary Digital Publishing Institute Proceedings*, 31(1), 84–91. <http://dx.doi.org/10.3390/proceedings2019031084>
- Yang, H., Kim, H., Shin, D., Lee, J. H. (2019). A study on adopting AI-based chatbot in elementary English-speaking classes. *Multimedia Assisted Language Learning*, 22(4), 184–205.
- Yin, J., Goh, T.-T., Yang, B., & Xiaobin, Y. (2020). Conversation technology with micro-learning: The impact of chatbot-based learning on students' learning motivation and performance. *Journal of Educational Computing Research*, 59(1). <https://doi.org/10.1177/0735633120952067>