

*G. Momonov<sup>1</sup>, L. Mirtskhulava<sup>2</sup>*

<sup>1</sup>Suleyman Demirel University, Kaskelen, Kazakhstan

<sup>2</sup>Tbilisi State University, Tbilisi, Georgia

## **ARTIFICIALLY INTELLIGENT CHATBOTS FOR HIGER EDUCATION: A REVIEW OF EMPIRICAL LITERATURE**

**Abstract.** This paper reviews empirical literature on artificially intelligent chatbots for higher education published in 2000-2021 in peerreviewed journals. The review shows that chatbots are mainly used for teaching assistance and for non-academic support purposes. The published studies unanimously report positive effects of chatbots on the outcomes of interest. The review suggests that chatbots have a considerable potential to be used more widely in higher education. Further, the research on chatbots in higher education is ripe for more research.

**Keywords:** chatbot, dialogue system, higher education, artificial intelligence, applications.

\*\*\*

**Аннотация.** В данной статье представлен обзор эмпирической литературы по чат-ботам с искусственным интеллектом для высшего образования, опубликованной в период 2000-2021 гг. в рецензируемых журналах. Обзор показывает, что чат-боты в основном используются для обучения и неакадемической поддержки. Опубликованные исследования единогласно сообщают о положительном влиянии чат-ботов на интересующие исследователей результаты. Обзор показывает, что чатботы имеют значительный потенциал для более широкого использования в высшем образовании. Кроме того, сфера изучения чат-ботов в высшем образовании требует дальнейших исследований.

**Ключевые слова:** chatbot, dialogue system, higher education, artificial intelligence, applications.

\*\*\*

**Андатпа.** Бұл мақалада 2000-2021 жылдары рецензияланған журналдарда жарияланған жоғары білімге арналған жасанды интеллектуалды чат-боттар туралы эмпирикалық әдебиеттер қарастырылады. Шолу көрсеткендей, чат-боттар негізінен оқытуда көмек көрсету үшін және академиялық емес қолдау мақсатында қолданылады. Жарияланған зерттеулер бірауыздан студенттердің нәтижелеріне

чатботтардың оң әсерін көрсетеді. Шолу чат -боттардың жоғары оқу орындарында кеңінен қолданылатын үлкен әлеуетке ие екендігін көрсетеді. Сонымен қатар, жоғары оқу орындарындағы чатботтар туралы зерттеулер саласы қосымша зерттеулерге дайын.

**Түйін сөздер:** чатбот, диалог жүйесі, жоғары білім, жасанды интеллект, қосымшалар.

### *1. Introduction*

Recent advances in artificial intelligence (AI) technologies boosted the widespread adoption of AI conversational chatbots in many areas of human activity. AI chatbots have an enormous potential to be fruitfully utilized in higher education due to their capacity to flexibly support large groups of students at a low cost [1-4]. Research on AI chatbots in higher education grew in the recent years yet is scattered across various disciplines and research methodologies which makes it difficult to systematically synthesize [5, 6].

Several reviews of the use of chatbots in education have been published recently. Winkler & Sollner [7] examine the applications of chatbots to improve learning processes and outcomes. Perez et al. [8] focus on chatbots' potential to improve learning and student services. Okonkwo and Ade-Ibijola [9] review chatbots' benefits and challenges in education. Wollny et al. [10] explore studies of chatbots focusing on applications, designs, evaluation, and educational effects.

While the reviews cited above synthesize an impressive amount of evidence, they tend to focus on secondary education and often adopt broad interpretations of chatbots. Arguably, higher education students have somewhat different needs than secondary education students, therefore, chatbot uses in higher education have their own unique features. Furthermore, most of the literature on chatbots in education describes various chatbots and their design features without any empirical research on chatbots' actual use. This review aims to systematically examine the burgeoning empirical literature on AI chatbots in higher education.

### *2. Review methodology*

#### *2.1. Research questions*

This review is guided by the following research questions: 1) What are chatbots used for in higher education? 2) How do chatbots affect academic and non-academic outcomes of students in higher education?

#### *2.2. Databases searched*

The search was conducted in the following databases relevant for computer science, education research, and social sciences: Association for Computing Machinery (ACM) digital library, Institute of Electrical and Electronics Engineers (IEEE), Computer Society Digital Library (CSDL), Education

Resources Information Center (ERIC), Scopus, and the National Bureau of Economic Research (NBER).

### 2.3. Search terms

Since chatbots are relatively recent and the literature on chatbots in higher education is in its emerging stage, the main search term we used was “chatbot”. To make sure the search is comprehensive, the search terms list was developed in consideration of possible studies that examine the use of chatbots to support a) students preparing to enter higher education, b) students already enrolled in institutions of higher education, and c) students preparing for the labor market or for subsequent stages of their educational trajectories. The search terms therefore included “chatbot for students”, “chatbot for education”, “chatbot for K-12”, “chatbot for university”, “chatbot for learning”, “chatbot for graduate education”, and “chatbot for the job market”.

### 2.4. Criteria for including papers in the review

The papers included in the review had to meet the following criteria: 1) demonstrate empirical evidence involving human participants on the use of chatbots in or for higher education, 2) be published in peer-reviewed journals, 3) be published in 2010-2021, 4) be published in English. After applying these inclusion criteria, our final analytic sample included 6 papers. *3. Results*

Table 1 presents an overview of the papers identified through the search and included in the final analytic sample.

*Table 1. Studies included in the review*

| <b>Study</b>             | <b>Research design</b> | <b>Country</b> | <b>Sample size</b> |
|--------------------------|------------------------|----------------|--------------------|
| Ayedoun et al. [11]      | Pre-/post-analysis     | Japan          | 5                  |
| Ciechanowski et al. [12] | RCT                    | Poland         | 31                 |
| Fryer et al. [13]        | Pre-/post-analysis     | Japan          | 122                |
| Lin & Chang [14]         | Mixed methods          | Canada         | 28                 |
| Page & Gehlbach [15]     | RCT                    | USA            | 7,489              |
| Nurshatayeva et al. [16] | RCT                    | USA            | 4,442              |

Ayedoun et al. [11] developed and evaluated a prototype of a chatbot encouraging willingness to communicate among university students studying English as a foreign language in Japan. Their chatbot prototype simulated a conversation at a restaurant. A total of 5 undergraduate and graduate students at a Japanese university interacted with the chatbot prototype. Before and after participants’ interaction with the chatbot, the researchers measured their willingness to communicate, interactions with the chatbot, nervousness, degree

of immersion, and fluency of conversation. The findings showed that confidence and desire for communication increased and nervousness decreased after interacting with the chatbot. Overall, the authors conclude that the chatbot increased participants' willingness to communicate.

Ciechanowski et al. [12] explored the psychophysiological reactions to chatbots and willingness to collaborate with a chatbot. Psychophysiological reactions were measured using electromyography, respirometer, electrocardiography, and electrodermal activity. Willingness to collaborate with chatbots was examined using the theory of planned behavior survey, the social presence survey, and the anthropomorphism scale. Two chatbots were designed for the Kozminski Academy to assist new students with enrollment related issues. One chatbot was text-based, the other chatbot had an avatar reading out the responses in addition to presenting them as text on screen. Participants (n=31) were randomly assigned to either of these chatbots. The overall conclusion of this study is that participants enjoyed their interactions with both chatbots. However, they were more positive about the text-based chatbot. Fryer et al. [13] examine the effects of using chatbots on task and course interest of students studying English as a foreign language at a private university in Japan. A total of 122 students from various majors were randomly assigned either to interacting with a chatbot or a human partner for one week. After the first week, the treatments were switched, that is, the chatbot group started chatting with a human and vice versa. The findings showed that the chatbot decreased students' task interest. Structural equation modeling showed that task interest with a human partner only contributed to increasing course interest. Notably, students' task interest was high at the start of interactions with the chatbot, but decreased sharply after the first task suggesting that chatbot's novelty effect faded quickly and did not contribute to sustained interest in learning.

Lin & Chang [14] present a chatbot aimed at improving university students' writing skills. Their mixed methods study with Canadian undergraduate students showed that the chatbot improved students' essay outline performance and was helpful in facilitating the communication between instructors and students.

Page & Gehlbach [15] conducted a randomized controlled trial with 7,489 students entering Georgia State University to estimate the effect of an AI chatbot on enrollment and academic outcomes in the first semester of university studies. The chatbot aimed to support students in transitioning from high school to college and enrolling in Georgia State University. The experiment results showed that the chatbot increased success with pre-enrollment tasks and raised enrollment by 3.3 percentage points among students who expressed early commitment to study at Georgia State University.

Nurshatayeva et al. [16] replicated the study by [15] at East Carolina University. A total of 4,442 students entering East Carolina University were randomly

assigned to either receiving the chatbot support or to business-as-usual supports provided by the university. The results suggest that the chatbot had no effect on overall enrollment rates but increased enrollment among first-generation students by 3 percentage points.

#### *4. Conclusion*

The present review showed that research papers evaluating chatbots in higher education focus on chatbots teaching English as a foreign language, developing writing skills, and supporting students in their transition from high school to university.

Overall, based on the reviewed studies, one can conclude that chatbots positively affect academic and non-academic outcomes of students in higher education. Chatbots assist students in improving their English language skills and other skills like writing. Chatbots also have the capacity to support students in the completion of various administrative tasks and nudge them to enroll on time in a university of their choice.

Notably, the reviewed papers tend to use rigorous research designs to evaluate the effects of the chatbots on the outcomes of interest. Randomized controlled trials were common perhaps reflecting the considerable research methods capacity of the researchers studying chatbots in education. Sample sizes in the two U.S.-based RCT studies are quite large reflecting the scale-up potential of chatbots in contrast to other types of educational technologies. Overall, the treatment effects of chatbots yield themselves particularly well to evaluations involving advanced quantitative research designs. The area of chatbots for higher education is ripe for more research.

### **References**

1. G. Momonov, "Artificially intelligent conversational chatbots: Main types, programming issues and possible solutions," *Proceedings of IYSW*, vol. 10, pp. 350-358, 2021.
2. A.-A. Georgescu, "Chatbots for education—trends, benefits and challenges," in *Conference proceedings of eLearning and Software for Education «(eLSE)*, 2018, vol. 2, no. 14: "Carol I" National Defence University Publishing House, pp. 195-200.
3. S. Yang and C. Evans, "Opportunities and challenges in using AI chatbots in higher education," in *Proceedings of the 2019 3rd International Conference on Education and E-Learning*, 2019, pp. 7983.
4. J. Wang, G.-H. Hwang, and C.-Y. Chang, "Directions of the 100 most cited chatbot-related human behavior research: A review of academic publications," *Computers and Education: Artificial Intelligence*, p. 100023, 2021.

5. I. Roll and R. Wylie, "Evolution and revolution in artificial intelligence in education," *International Journal of Artificial Intelligence in Education*, vol. 26, no. 2, pp. 582-599, 2016.
6. T. Araujo, "Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions," *Computers in Human Behavior*, vol. 85, pp. 183-189, 2018.
7. R. Winkler and M. Söllner, "Unleashing the potential of chatbots in education: A state-of-the-art analysis," 2018.
8. J. Q. Pérez, T. Daradoumis, and J. M. M. Puig, "Rediscovering the use of chatbots in education: A systematic literature review," *Computer Applications in Engineering Education*, vol. 28, no. 6, pp. 1549-1565, 2020.
9. C. W. Okonkwo and A. Ade-Ibijola, "Chatbots applications in education: A systematic review," *Computers and Education: Artificial Intelligence*, p. 100033, 2021.
10. S. Wollny, J. Schneider, D. Di Mitri, J. Weidlich, M. Rittberger, and H. Drachsler, "Are we there yet?-A systematic literature review on chatbots in education," *Frontiers in artificial intelligence*, vol. 4, 2021.
11. E. Ayedoun, Y. Hayashi, and K. Seta, "A conversational agent to encourage willingness to communicate in the context of English as a foreign language," *Procedia Computer Science*, vol. 60, pp. 1433-1442, 2015.
12. L. Ciechanowski, A. Przegalinska, M. Magnuski, and P. Gloor, "In the shades of the uncanny valley: An experimental study of human– chatbot interaction," *Future Generation Computer Systems*, vol. 92, pp. 539-548, 2019.
13. L. K. Fryer, M. Ainley, A. Thompson, A. Gibson, and Z. Sherlock, "Stimulating and sustaining interest in a language course: An experimental comparison of Chatbot and Human task partners," *Computers in Human Behavior*, vol. 75, pp. 461-468, 2017.
14. M. P.-C. Lin and D. Chang, "Enhancing Post-secondary Writers' Writing Skills with a Chatbot," *Journal of Educational Technology & Society*, vol. 23, no. 1, pp. 78-92, 2020.
15. L. C. Page and H. Gehlbach, "How an artificially intelligent virtual assistant helps students navigate the road to college," *AERA Open*, vol. 3, no. 4, pp. 1-12, 2017.
16. A. Nurshatayeva, L. C. Page, C. C. White, and H. Gehlbach, "Are artificially intelligent conversational chatbots uniformly effective in reducing summer melt? Evidence from a randomized controlled trial," *Research in Higher Education*, vol. 62, no. 3, pp. 392-402, 2021.