

Effect of Guidance-Based ChatGPT Learning Aid (GCLA) on College Students' Reading Comprehension: An Experimental Study

Adin Fauzi¹, Muhammad Iqbal Baihaqi², Linda Alif Dinata³

^{1, 2, 3}Faculty of Teacher Training and Education, Universitas Islam Balitar, Indonesia

*) correspondence: adinfauzi2693@gmail.com

ABSTRACT

The rapid advancement of technology demands that educators continually innovate in their teaching methods. ChatGPT, a cutting-edge artificial intelligence (AI) technology, holds immense potential for English pedagogy. This study aims to investigate the impact of the Guidance-based ChatGPT Learning Aid (GCLA) on students' reading comprehension. This pre-experimental study employed a one-group pretest-posttest design, involving twenty-one students from the Primary School Teacher Education Study Program. The students participated in a pretest, received treatments in four weeks, and then took a posttest. The main data collection instrument was tests, and the resulting data were analyzed using the Wilcoxon signed-rank test due to the fact that the data from the pretest and posttest were not in a normal distribution. The findings showed that after four weeks, GCLA produced a statistically significant improvement in students' reading comprehension. The result of the Wilcoxon signed-rank test revealed a p-value of < 0.001 , which is less than the critical value of 0.05, resulting in the rejection of the null hypothesis. This indicates that using ChatGPT in reading instruction, along with the guidance-based approach, is able to augment the students' ability to grasp the ideas of the given passages. This study suggests that AI technology, when used properly, can have a positive impact on both the learning process and student achievement.

Keywords: ChatGPT; guidance; GCLA; reading comprehension

INTRODUCTION

Reading is crucial for students as it significantly contributes to a smooth learning process. It not only provides valuable information but also makes learning enjoyable, which is particularly important for foreign language acquisition (Failasofah, 2014). When it comes to learning English, strong reading skills can enhance vocabulary, improve reading fluency, boost text comprehension, and develop critical thinking abilities (Anaktototy & Lesnussa, 2022; Lalić & Dubravac, 2021).

In the context of higher education, numerous studies have identified key difficulties in reading faced by college students. Anwar and Sailuddin (2022) revealed several problems in academic reading, such as specialized vocabulary, the use of language, ideas of the passages, the length of the texts, reading skills, and challenging topics. In addition, Isma and Nur (2023) found that the majority of students struggled with understanding English passages. They further stressed the importance of effective reading strategies and the integration of learning media and technological tools to address those challenges.

One effective tool that has recently been used for teaching and learning English is ChatGPT. Developed by OpenAI, ChatGPT is a large language model praised for its remarkable proficiency in generating seemingly precise responses in various languages, and it is particularly adept at dialog, excelling in human-like interactions (Liu & Ma, 2023; De Angelis, 2023). As a leading generative Artificial Intelligence (AI) chatbot, ChatGPT has the potential to transform the methods used for learning, teaching, and assessing language (Bhullar et al., 2024; Liu & Ma, 2023). Several studies have pointed out this issue. For instance, Park (2023) highlighted its ability to provide dynamic feedback and personalized learning experiences. Moreover, Kostka and Toncelly (2023) suggested that ChatGPT is a beneficial resource in English Language Teaching (ELT) as it offers learners numerous opportunities to practice and enhance their skills. In reading practice, educators can utilize ChatGPT

to develop reading tasks or to explain challenging texts. ChatGPT's capacity to condense or elucidate difficult passages can improve students' comprehension. Furthermore, Bin-Hady et al. (2023) revealed that ChatGPT can enhance students' language skills by supporting the learning process, providing feedback on language use, and acting as a partner for language practice with recommended activities. Not to mention, Liu and Ma (2023) found that learners with positive attitudes towards ChatGPT's usefulness tend to use it for informal English learning.

Additionally, an interesting research by Lee et al. (2024) investigates the implementation of the Guidance-based ChatGPT Learning Aid (GCLA). The GCLA marks a significant innovation in learning process, enhancing the interaction between learners and technology. By leveraging ChatGPT's extensive knowledge, GCLA effectively responds to a wide range of student questions. Instead of providing direct answers, GCLA encourages students to develop their own responses, thereby promoting deep engagement with the material and fostering higher-order thinking skills (HOTS). In their study, GCLA was found to have positive and significant effects on students' cognitive abilities, engagement, self-efficacy, critical thinking, problem-solving, creativity, and knowledge construction. Their study, which focused on chemistry learning, indicates that further research is necessary to explore the impact of GCLA in various domains and contexts. Although we used the same terminology for the approach, we made adjustments based on available infrastructure. As they explained in their study, they used Apple's Shortcuts, a user-friendly platform for custom task automation, alongside ChatGPT. While our study did not exactly employ the same tool, we adopted a similar process to examine its effect on English learning.

The present study examines the application of GCLA in reading instructions. In particular, it aims to investigate the impact of ChatGPT on students' reading comprehension. The findings are expected to offer valuable insights into how GCLA and ChatGPT can be effectively utilized in the English language learning process, particularly in higher education.

METHODS

Twenty-one students from the Primary School Teacher Education Study Program participated in this research. Besides improving their English language skills, this study specifically aims to prepare these students to teach English in elementary schools once they become teachers.

Furthermore, due to the small number of students, the researchers were unable to use two separate classes as experimental and control groups. Dividing the students into two groups was also impractical, as it would disrupt the lecture system. Therefore, the researchers adopted a one-group pretest-posttest design (Ary et al., 2010). The students first took a pretest, then received the treatment in four weeks, and finally took a posttest. The comparison of pretest and posttest results was used to evaluate the impact of the treatment.

The primary tool for collecting data in this study was tests. The researchers used commercially available test instruments from EnglishScore, which have been validated and proven reliable. To confirm this, researchers reviewed the Validity Report published by EnglishScore (British Council EnglishScore, 2023). In accordance with the EnglishScore test format, the researchers used multiple-choice questions with the focus of analyzing the gist and details of the given passages. The test's level was intermediate, ranging from B1 to B2 based on the CEFR levels (Council of Europe, 2020). This level was chosen based on the result of a diagnostic test run at the beginning of the semester, showing that the students' reading ability was somewhere between beginner and intermediate. To achieve improvement, the researchers decided to use intermediate levels both in the pretest and posttest. The scores for both the tests were based on the percentage of correct answers.

The data from the tests were analyzed using the Wilcoxon signed-rank test due to the fact that the data from the pretest and posttest were not in a normal distribution. This nonparametric test was conducted to evaluate the impact of GCLA on students' reading comprehension. For drawing conclusions, the null hypothesis is rejected if the test statistic is less than 0.05 as the critical value (King & Eckersley, 2019). In this study, the null hypothesis states: "The application of GCLA does not have a significant effect on students' reading comprehension."

RESULTS AND DISCUSSION

This study seeks to answer a research question: “Does GCLA produce significant improvement in college students’ reading comprehension?” To answer this question, the researchers conducted a series of activities, which included a pretest, treatments, and posttest. In both the pretest and posttest, the researchers used an online quiz platform. Both tests were conducted in the classroom, in which the reading activities were controlled by the researchers. In doing this, the students answered each question, and once all the students finished answering, they moved on to the next ones. A three-minute time limit was given for the students to answer each question. The topic of the questions relates to their field of study, education, especially the teaching profession. This topic was chosen by considering the intermediate level as described in the CEFR level; the students can comprehend nonfiction texts discussing specific materials relating to their field of study (Council of Europe, 2020). In this study, comprehension is defined as the students’ ability to obtain information and ideas from the given passages.

Between the pretest and posttest, the students followed a four-week treatment, which consisted of one meeting per week lasting about 100 minutes. These treatments focused on training them to understand the ideas from the passages with the help of ChatGPT. In giving the treatments, the researchers provided different texts each week. The main classroom activities covered reading the texts, generating questions, providing initial answers, and using ChatGPT to confirm the answers.

The treatment began with the provision of various texts to the students. These texts were carefully selected to ensure broad exposure to contents relating to education. The goal was to engage students with materials that were both interesting and challenging. This could stimulate their curiosity and encourage deeper engagement with the content.

Prior to the distribution of texts, a discussion was initiated to explore the students' background knowledge. This step was crucial in identifying what the students already knew about the topics presented in the texts. By discussing their pre-existing knowledge, students could build a foundation for new learning and connect new information to what they already understood. This step also helped in identifying keywords that could possibly construct the texts.

After distributing the texts, students read the texts, and they were guided to generate their own questions about the texts. This phase encouraged them to think critically and analytically about the contents they had read. Generating questions helped students actively engage with the content, promoting a deeper understanding.

Once the students had formulated their questions, they were asked to provide initial answers. These answers could be either correct or incorrect, as the purpose was to capture their initial thoughts and reasoning processes. This step allowed students to articulate their understanding and provided a basis for further exploration and discussion.

To facilitate the reading process, a model chat (prompt) of ChatGPT was provided. The prompt was designed to guide students in discussing their initial answers, empowering their critical thinking. ChatGPT played a role in analyzing the students' responses, offering suggestions and insights without providing immediate answers. This approach encouraged students to reflect on their answers and consider different perspectives, promoting a deeper engagement with the material. The following figures show each prompt and the corresponding ChatGPT conversations:

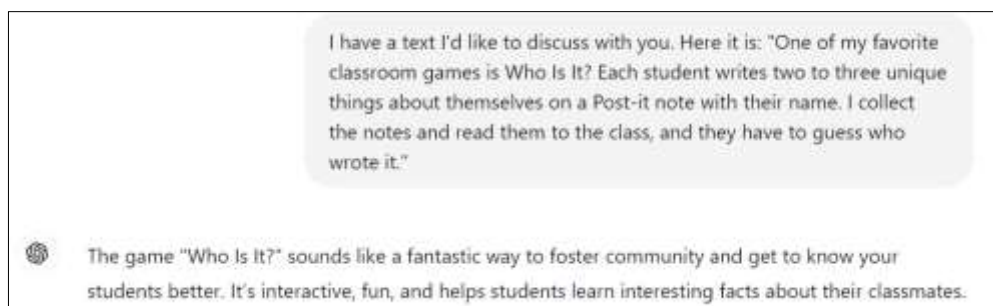


Figure 1. Example of prompt 1 and its output

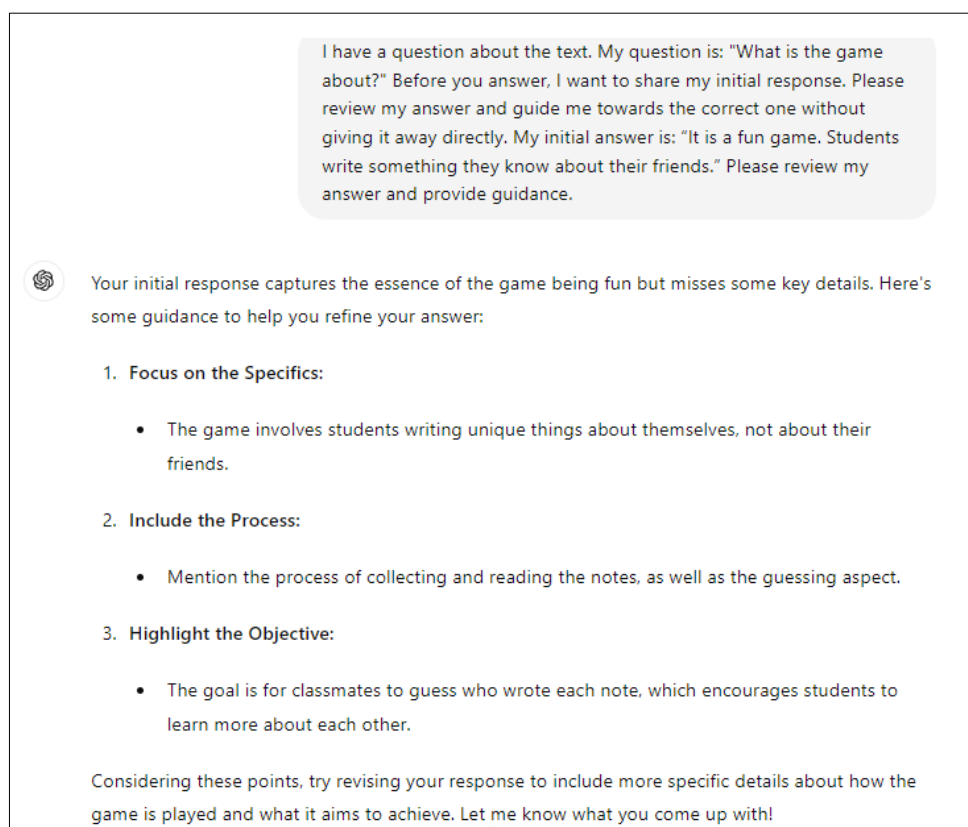


Figure 2. Example of prompt 2 and its output

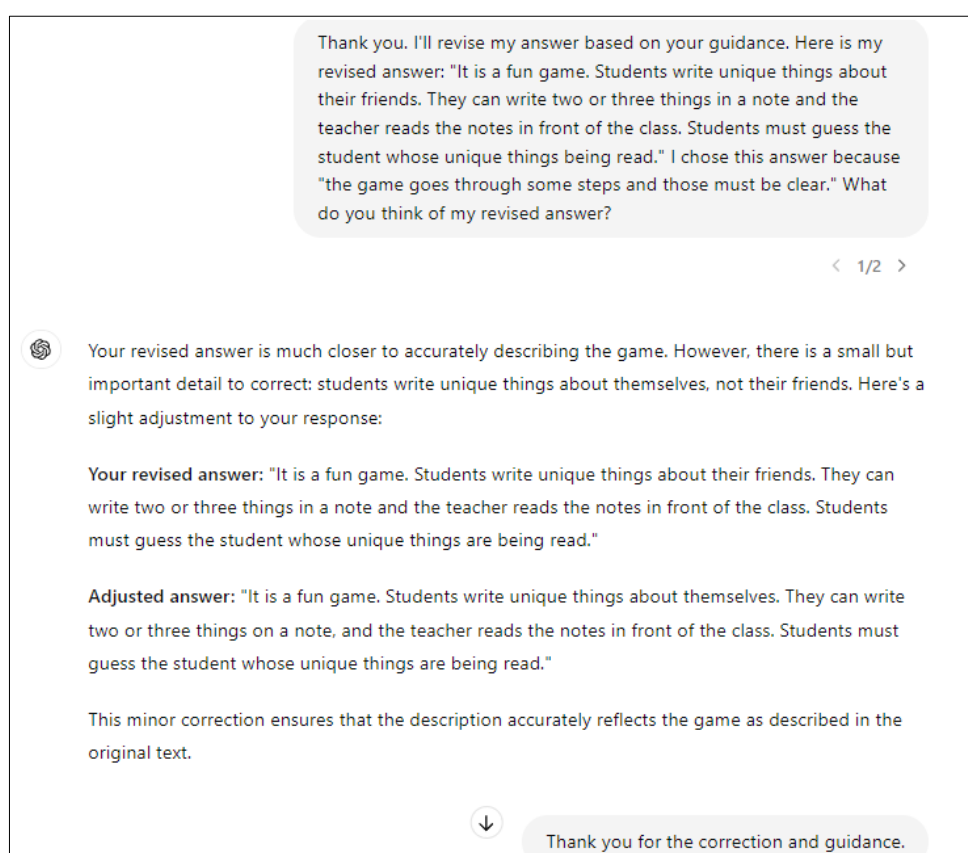


Figure 3. Example of prompt 3 and its output

Along with the use of ChatGPT's prompts above, the students were supervised to ensure they eventually arrived at the correct answers. The supervision involved providing guidance and support as needed, helping students navigate the discussion and refine their understanding. The use of ChatGPT as a discussion room rather than an answer provider was intended to train students to use AI tools wisely, maximizing their own thinking and problem-solving abilities. This type of practice was carried out in four meetings, with each week featuring a unique text. Although the texts varied, the topics remained consistent, focusing on education, particularly the teaching profession.

After going through the treatments, the students took the posttest. The result of the posttest revealed meaningful progress in student comprehension. From the descriptive analysis, the mean of the posttest surpassed the mean of the pretest. Figure 4 shows the results of the descriptive statistics.

Descriptive Statistics					
	N	Mean	Std. Deviation	Minimum	Maximum
Pretest	21	53.33	16.803	18	91
Posttest	21	79.62	17.755	40	100

Figure 4. The result of descriptive statistics

As shown in Figure 4, the mean of posttest score (79.62) is higher than the mean of pretest score (53.33). Additionally, the minimum score in the posttest (40) surpasses the pretest minimum score (18). The maximum score also improved, rising from 91 in the pretest to 100 in the posttest. These results indicate a significant improvement in students' reading comprehension.

In addition to the previous analysis, the researchers also examined the rank output from the Wilcoxon-signed rank test. This output provides valuable information about the negative ranks (a decrease in posttest scores), positive ranks (an increase in posttest scores), and ties (no change between pretest and posttest scores). Figure 5 shows the rank output.

Ranks				
		N	Mean Rank	Sum of Ranks
Posttest - Pretest	Negative Ranks	0 ^a	.00	.00
	Positive Ranks	21 ^b	11.00	231.00
	Ties	0 ^c		
	Total	21		

a. Posttest < Pretest
b. Posttest > Pretest
c. Posttest = Pretest

Figure 5: The rank output

As shown in Figure 5, the negative ranks indicate that no student scored lower on the posttest compared to the pretest. The positive ranks demonstrate that all students improved their scores on the posttest. Regarding ties, the data reveals that none of the students had identical scores on both the pretest and posttest. In other words, every student showed improvement in their reading scores, with an average increase (the mean rank) of 11 points.

Finally, to test the hypothesis, the researchers analyzed the test statistics from the Wilcoxon signed-rank test. We examined the Asymp. Sig. (2-tailed) value to determine whether to reject the null hypothesis. If this value is less than 0.05, the null hypothesis is rejected. Figure 6 displays the test statistics output.

Test Statistics ^a	
	Posttest - Pretest
Z	-4.016 ^b
Asymp. Sig. (2-tailed)	<.001

a. Wilcoxon Signed Ranks Test

b. Based on negative ranks.

Figure 6. The test statistics output

Based on the output above, the Asymp. Sig. (2-tailed) value is < 0.001 , which is less than the critical value of 0.05. This statistical result indicates that the four-week treatment led to a change in students' abilities. Therefore, there is sufficient evidence that implementing GCLA significantly improved students' reading comprehension.

Reflecting on the results, this study highlights several key findings. Firstly, similar to the findings of Bhullar et al. (2023), the application of GCLA has the potential to revolutionize the learning process. By integrating technology, GCLA creates an engaging environment where students are active participants in the learning process. The use of technology in GCLA fosters a dynamic and interactive classroom setting. In the context of reading activities, students had more opportunities to engage with the texts. They were enthusiastic about reading the texts, posing questions, providing initial answers, understanding the output from ChatGPT, reading its explanations, reflecting on their own answers, and revising their answers with guidance from ChatGPT. These findings support Kostka and Toncelly's (2023) claim that ChatGPT can enhance students' engagement with materials and make the process of examining ChatGPT's output enjoyable. Consequently, it can be said that ChatGPT is an effective tool for students to use in digitalized formal language learning environments (Liu & Ma, 2023; Park, 2023).

Secondly, the guidance provided by GCLA played a crucial role in enhancing students' reading abilities. Pre-class discussions combined with ChatGPT prompts and outputs significantly improved students' comprehension. These pre-class activities prepared students by activating their prior knowledge. When ChatGPT was used, students were guided to explore the texts more thoroughly. They first generated questions under the supervision of the researchers, focusing on specific words or parts of the texts. After providing their initial answers, they used ChatGPT to check their responses with the given prompts. This process led to better understanding, as students reflected on ChatGPT's feedback and made revisions until they arrived at the correct answers. This finding supports Lee et al.'s (2024) study, demonstrating that GCLA, rather than providing direct answers, helps students improve their achievement through guidance. In addition, the effectiveness of guidance in GCLA for improving comprehension aligns with findings from a study by Anaktototy and Lesnussa (2022). Their research showed that directed reading and thinking activities, which involve guiding students through the reading process, can enhance both reading abilities and critical thinking. Both their study and ours suggest that engaging with the text through predicting and validating answers leads to more interactive and effective reading lessons.

All the practices in GCLA are instrumental in training students to think independently and critically, reducing their dependence on technology. With consistent practice, the students became adept at applying reading strategies, such as using context clues and leveraging their existing knowledge and vocabulary. This approach not only enhanced their comprehension skills but also prepared them to tackle reading tasks with confidence and competence, especially during tests.

CONCLUSION

This study found that implementing GCLA led to significant improvements in students' reading comprehension. The reading process, which involved deep engagement with the texts through questioning, predicting, and validating answers, combined with the use of ChatGPT, enhanced the

students' ability to comprehend the passages. This study demonstrates that AI technology, despite educators' concerns about academic integrity, can be used safely and effectively with proper guidance. The critical thinking involved in the reading process helped students improve their reading skills and become accustomed to the thinking process useful for reading tests. Although this study provides insights into teaching reading in higher education, it has limitations, including a small number of participants and the absence of a control group. Future research could involve more participants and include a control group in a similar experiment to provide more detailed information about the effects of GCLA on reading comprehension. Additionally, similar studies could be conducted in other areas of English pedagogy, such as listening, speaking, writing, vocabulary, and grammar. This would enrich the evidence on how to effectively integrate AI technology into English teaching and learning.

ACKNOWLEDGEMENTS

The researchers would like to thank the Department of Research and Community Service (LPPM) at Universitas Islam Balitar for their financial support of this research. Additionally, sincere appreciation goes to the cooperative students for their willingness to assist in this study.

REFERENCES

- Anaktototy, K., & Lesnussa, I. (2022). Improving EFL students' reading comprehension and critical thinking skills through directed reading thinking activity. *Eralingua*, 6(1), 244-255. <https://doi.org/10.26858/eralingua.v6i1.27711>
- Anwar, I. W., & Sailuddin, S. P. (2022). Academic reading difficulties in higher education. *JOLLT Journal of Languages and Language Teaching*, 10(2), 309-314. <https://doi.org/10.33394/jollt.v%vi%oi.4849>
- Ary, D., Jacobs, L. C., Sorensen, C., & Razavieh, A. (2010). *Introduction to research in education*. Wadsworth: Cengage Learning.
- Bhullar, P. S., Joshi, M., & Chugh, R. (2024). ChatGPT in higher education - a synthesis of the literature and a future research agenda. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-024-12723-x>
- Bin-Hady, W. R. A., Al-Kadi, A., Hazaea, A., & Ali, J. K. M. (2023). Exploring the dimensions of ChatGPT in English language learning: A global perspective. *Library Hi Tech*. <https://doi.org/10.1108/LHT-05-2023-0200>
- British Council EnglishScore. (2023). *Core skills test: Purpose and content*.
- Council of Europe (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*. Council of Europe Publishing, Strasbourg.
- De Angelis, L., Baglivo, F., Arzilli, G., Privitera, G. P., Ferragina, P., Tozzi, A. E., & Rizzo, C. (2023). ChatGPT and the rise of large language models: the new AI-driven infodemic threat in public health. *Frontiers in Public Health*, 11, 1-8. <https://doi.org/10.3389/fpubh.2023.1166120>
- Failasofah. (2014). Practices in generating elementary school students' literacy set. *LearnING Journal*, 1(1), 17-39.
- Isma, A., & Nur, S. (2023). Examining the reading performance and problems among students in Indonesian higher education. *Scope of English Language Teaching, Literature and Linguistics*, 6(1), 12–23. <https://ejournals.umma.ac.id/index.php/seltics>
- King, A. P., & Eckersley, R. J. (2019). Inferential statistics III: Nonparametric hypothesis testing. In A. P. King & R. J. Eckersley (Eds.), *Statistics for biomedical engineers and scientists* (pp. 119-145). Academic Press. <https://doi.org/10.1016/B978-0-08-102939-8.00015-3>
- Kostka, I., & Toncelli, R. (2023). Exploring applications of ChatGPT to English language teaching: Opportunities, challenges, and recommendations. *TESL-EJ*, 27(3), 1-19. <https://doi.org/10.55593/ej.27107int>
- Laličić, A., & Dubravac, V. (2021). The role of reading in English language classrooms. *MAP Social Sciences*, 1(1), 23–36. <https://doi.org/10.53880/2744-2454.2021.1.1.23>

- Lee, H. Y., Chen, P. H., Wang, W. S., & et al. (2024). Empowering ChatGPT with a guidance mechanism in blended learning: Effect of self-regulated learning, higher-order thinking skills, and knowledge construction. *International Journal of Educational Technology in Higher Education*, 21(16). <https://doi.org/10.1186/s41239-024-00447-4>
- Liu, G., & Ma, C. (2023). Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model. *Innovation in Language Learning and Teaching*, 18(2), 125–138. <https://doi.org/10.1080/17501229.2023.2240316>
- Park, H.-Y. (2023). Application of ChatGPT for an English learning platform. *STEM Journal*, 24(3), 30–48. <https://doi.org/10.16875/stem.2023.24.3.30>