Journal of World Englishes and Educational Practices
ISSN: 2707-7586
DOI: 10.32996/jweep
Journal Homepage: www.al-kindipublisher.com/index.php/jweep

JWEEP
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

ChatGPT in Language Education: Meta-Analysis

Muhammad Iqbal Amrullah[1], Firmansyah[2], Anisah Nabilah[3], Hendriwanto[4] ✉ and Ratna Muntiningsih[5]
[123]English Education Department, Swadaya Gunung Jati University, Cirebon, Indonesia
[45]English Education Lecturer, English Education Department, Swadaya Gunung Jati University, Cirebon, Indonesia
Corresponding Author: Hendriwanto, M.Pd, **E-mail**: hendriwanto46@gmail.com

| ABSTRACT
This research aims to explore the potential of ChatGPT in the educational context, including how this technology can enhance the learning and teaching experience. The data were searched electronically and via reference-list checking. For electronic searches, the main databases were SpringerLink, a Web of Science (WOS), and IEEE Xplore. The method employed is meta-analysis, which involves gathering and thoroughly analyzing data from multiple relevant scientific studies to recognize common patterns or trends and assess the cumulative impact of identical variables examined across each study. The results also suggest that in terms of English language abilities, writing is the most frequently utilized skill, and the predominant issue highlighted in studies is that of student integrity. This article aims to provide a comprehensive analysis of research on ChatGPT to initiate an evidence-based discussion on the usage of ChatGPT in English language education.

| KEYWORDS

ChatGPT, Artificial Intelligence, Language Learning, English Language Education.

| ARTICLE INFORMATION

ACCEPTED: 01 June 2024 **PUBLISHED:** 01 June 2024 **DOI:** 10.32996/jweep.2024.6.2.2

1. Introduction

In recent years, educational researchers have invested in the emerging issue of the use of artificial intelligence in education as a tool to prompt the writing section (2022–2023). For example, the study by Guo and Wang (2023) studied ChatGPT's significant contribution to teacher feedback on Chinese students' writing in the setting of second language writing. Another study conducted by F. Jia et al. (2022) examined how AI technologies can help with the development of sophisticated learning, including language learning. ChatGPT has been proven to be a platform for language instruction using the Neouron language program (NLP) developed by OpenAI, which uses enormous data sets to create text responses to inquiries, comments, and requests from students (Gilson et al., 2023). The previous study showed that ChatGPT contributed to the input of language learning for EFL students in the context of Asian countries (Jeon & Lee, 2023; Kohnke et al., 2023; Tlili et al., 2023).

As an example, learners have the capability to utilize ChatGPT to access dictionary definitions and illustrations (Kohnke et al., 2023, p. 3). ChatGPT can also improve the teaching experience for educators by acting as a conversational partner, a source of content, an assistant in teaching tasks, and an evaluator (Jeon & Lee, 2023, p. 9). Additionally, ChatGPT can improve educational teaching and learning results by serving as a provider of educational content. It offers teachers and students foundational knowledge in a wide range of subject areas using clear and straightforward language (Kasneci et al., 2023; Tlili et al., 2023).

Within the realm of artificial intelligence, natural language processing (NLP) has resulted in the development of cutting-edge virtual aids and chatbots capable of comprehending and producing human language (Farrokhnia et al., 2023). Holmes and Tuomi (2022) provided an instance of students transitioning to digital learning platforms where artificial intelligence is integrated to tailor and

Copyright: © 2024 the Author(s). This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) 4.0 license (https://creativecommons.org/licenses/by/4.0/). Published by Al-Kindi Centre for Research and Development, London, United Kingdom.

support students in personalizing their educational journeys. Presently, students enjoy greater accessibility to writing tools powered by artificial intelligence, offering them instant feedback and correction features for grammar, punctuation, and expression.

To address this gap, we empirically demonstrate in this study that AI In this meta-analysis study, we explore the ChatGPT in three areas of applied linguistics and TESOL, such as a) pedagogy, b) assessment, and c) competence.

2. Literature Review

ChatGPT can be defined as a generative AI with the potential to shift paradigms in education due to its considered pedagogical capabilities and benefits. (Jeon & Lee, 2023; Kohnke et al., 2023; Tlili et al., 2023). Other studies by (Coniam, 2008; Jeon & Lee 2023; Kohnke et al., 2023) have demonstrated that AI technologies can enable the implementation of teaching-centered strategies, leading to more dynamic, engaging, and sophisticated classroom pedagogies.

In simpler terms, AI technologies like ChatGPT hold promise for significantly enhancing education by offering valuable support alongside various teaching methods, such as scaffolding techniques and creating educational activities (Lo, 2023; Rahman & Watanobe, 2023). Also, artificial intelligence technology like ChatGPT has the potential to transform paradigms in education by enabling the implementation of more dynamic and sophisticated teaching strategies and providing valuable support for various teaching methods.

The previous study showed that ChatGPT contributed significantly to language teaching and learning, especially in the context of AI in language pedagogy, as our findings can be used to inform some of the established pedagogical-technological models and frameworks in the field. (Bedoya et al., 2023).

Research in higher education on assessment and evaluation has addressed various topics such as e-assessment (Bearman, Nieminen, and Ajjawi 2023), online exams (Dawson, Nicola-Richmond, and Partridge 2023), academic integrity concerns in online exams (Huber et al. 2023), and teachers' perspectives on e-assessments (Mimirinis 2019). This suggests that further investigation and attention are needed to understand the effects of technological advancements on assessment methodologies. Currently, there has been limited examination of issues related to generative artificial intelligence (GAI) and the impact of AI chatbots on assessment practices. For example, a study evaluating ChatGPT's performance in law school exams by analyzing its responses to four genuine exams found that ChatGPT achieved an average grade of C+ (Choi et al., 2023). In another study, ChatGPT was assessed in various open-ended questions from an MBA course in operations management, resulting in responses scoring between B and B- (Terwiesch, 2023). Liebrenz et al. (2023) highlighted the heightened severity of this ethical concern due to ChatGPT's susceptibility to producing inaccurate responses, thereby amplifying the potential dissemination of misinformation in scholarly publications. This could explain why certain studies offer limited examinations of student anxieties in academic essay writing, as prior reviews have already provided comprehensive insights into student concerns. Nonetheless, there are potentially additional aspects to investigate. Future researchers could opt to delve more deeply into these concerns to refine their focus (Farhi F et al., 2023). Gorichanaz T. (2023) also investigated students' reactions to allegations of utilizing ChatGPT in composing essays and other academic evaluations while also exploring the repercussions of AI technology integration in higher education and seeking suitable policy and educational remedies. While the thematic examination in this study yields valuable insights, it may not encompass all the intricate facets of students' encounters and reactions to accusations of ChatGPT-related cheating. Employing additional analytical techniques, such as quantitative analysis or in-depth interviews, could provide a more comprehensive understanding.

Zhuo et al. (2023) reported that although ChatGPT exhibits slightly superior performance compared to other current AI tools, it still demonstrates indications of ethical concerns, particularly in perpetuating social stereotypes and fostering unfair discrimination. Likewise, in research involving three seasoned university educators who utilized ChatGPT for an entire week, the educators reported that ChatGPT is susceptible to inaccuracies, including the introduction of bias (Tlili et al., 2023). Similarly, Newton and Xiromeriti (2023) found that ChatGPT underperformed compared to the average student in various fields, such as ophthalmology, law, economics, and physics, in multiple-choice tests.

3. Methodology

Electronic journal articles published between 2021 and 2024 were systematically searched using databases such as SpringerLink, Web of Science (WOS), and IEEE Xplore. These databases were selected for their coverage of high-impact journals indexed in the Science Citation Index and the Social Sciences Citation Index. Specific keywords related to the research objectives were utilized, including ChatGPT and English language learning. Additionally, reference-list checking was conducted to ensure comprehensive coverage of relevant literature.

3.1 Article selection

This study used many inclusion and exclusion criteria to restrict the number of search results for a reliable review.

3.1.1 First selection

In general, titles and abstracts were reviewed to select papers that met the following criteria: (a) published between January 2021 and December 2024; (b) written in the English language; (c) full-text version available; and (d) empirical research. As the search tools were different for each database, the number of retrieved articles was sometimes too large and unreadable. Therefore, different filtering techniques were applied between the databases. For example, when the keywords "mobile learning" and "English language" were used, too many research papers were retrieved. In SpringerLink, 9226 papers were retrieved. To reduce this number, the search results were filtered in the following manner: Content Type: Article; Discipline: Computer Science, Education, and Language. This resulted in 95 papers. In WOS, the initial search results of 182 papers were reduced to a shortlist of 78 when "article" was selected as the document type. Different search strings were not applied on IEEE Xplore. The final number of search results gathered using the above search strategies was 171 papers (SpringerLink 32, WOS 77, IEEE 62).

3.1.2 Second selection

During the second stage, each paper's title was reviewed and analyzed to determine whether it contained empirical evidence relating to the use of technology to solve problems in English language learning, especially limitations, challenges, and/or the use of mobile learning. The number of papers at the end of this stage was 107.

3.1.3 Third selection

In this stage, three experts, each with several years of experience in research on mobile and distance learning, were invited to confirm the relevance of all selected papers. The final number of papers at the end of this stage was 24. The total number of selected papers was 30. Two papers were repeated, and hence, the actual number of papers reviewed was 29, as shown in Table 2.

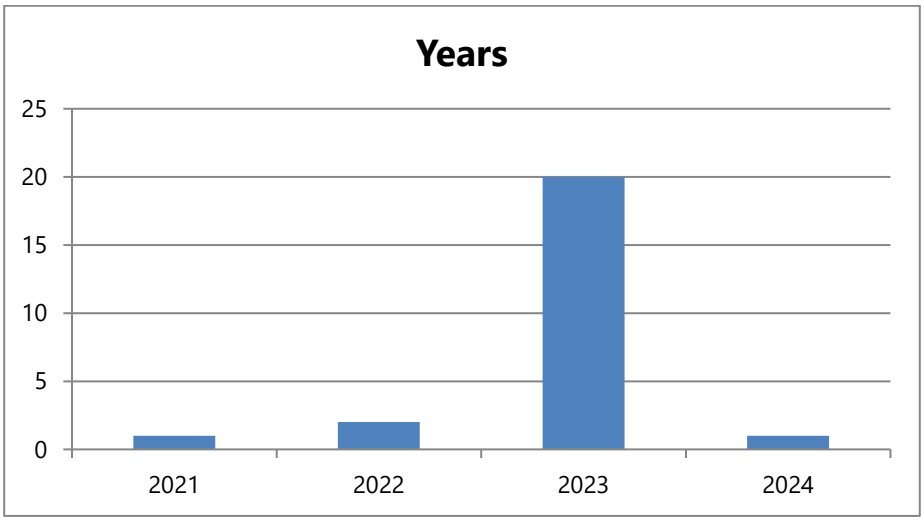
4. Results and Discussion

Table 1. Geographical location

Country	Number
Australia	1
China	3
Germany	2
Hongkong	2
Macau	1
Netherlands	3
Philippines	1
Poland	1
Singapore	4
Sweden	1
Thailand	1
United Arab Emirates	1
United States of America	2
United Kingdom	2

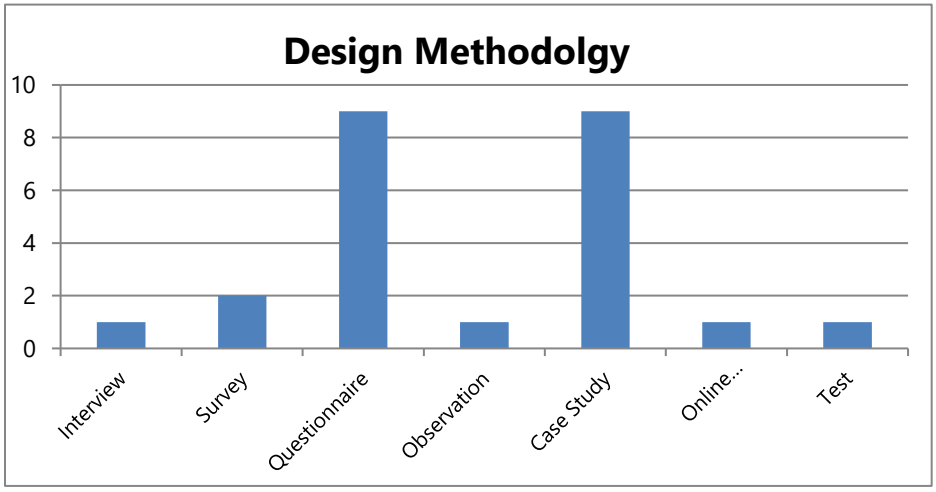
As shown in Table 4, of the 24 contributing countries, Singapore has published the most studies (4 articles), with China and the Netherlands second (3 articles). There are two publications from Germany, Hongkong, the United States of America, and the United Kingdom. Other countries produce one publication respectively (Australia, Macau, Philippines, Poland, Sweden, Thailand, United Arab Emirates).

Table 2. Year of Publications.



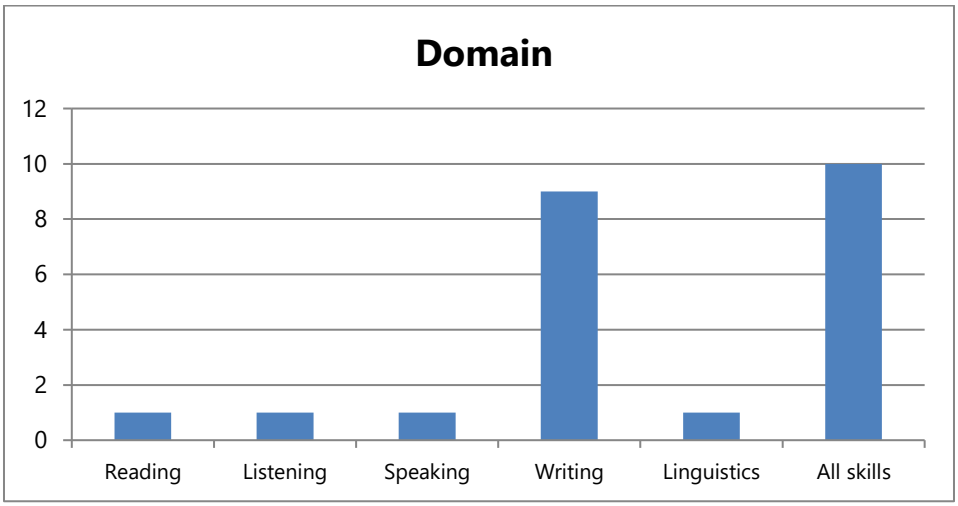
In 2023, the distribution of research papers is the highest number of publications. The results show that the number of publications at WOS and Scopus has increased over the years, with a few drops; however, SpringerLink does not follow this trend as well as Taylor and Francis.

Table 3. Design Methodology.



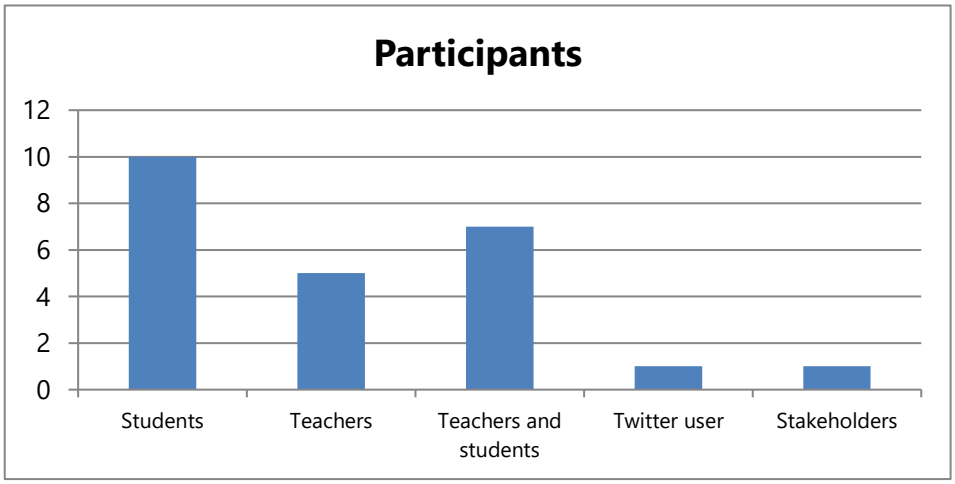
The design of the studies varies from one paper to another. Most designs use surveys and questionnaires, while other publication uses one method respectively (interview, observation, online discussion, and test).

Table 4. Research Domain.



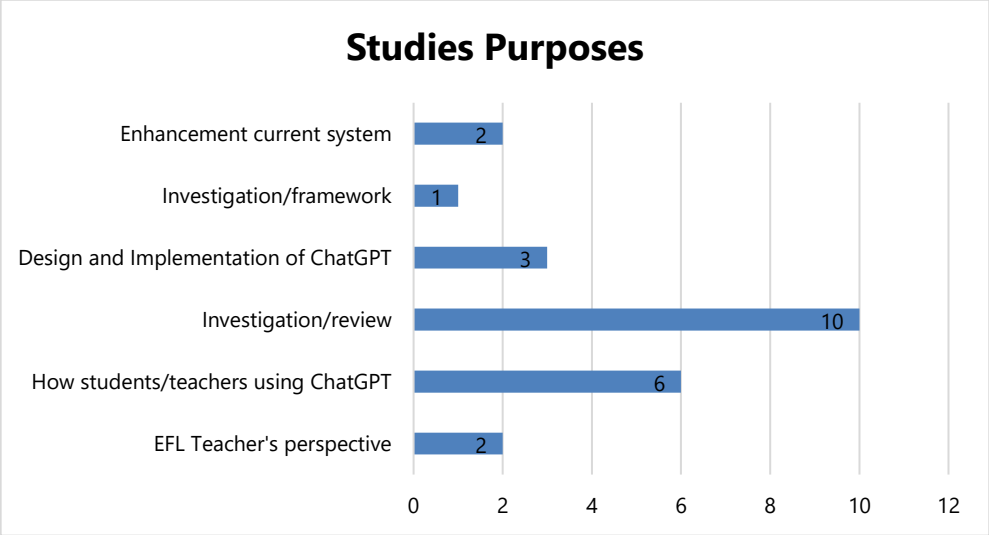
As shown in Table 4, there are 10 papers focused on improving all the English language skills or only mentioned "English language" without providing more details. Followed by 9 papers that focused on writing skills. While other papers focus on one domain respectively (reading, listening, speaking, and linguistics).

Tabel 5. Participants.



As shown in Figure 4, most participants consisted of students (10 papers), followed by teachers and students (7 papers), and 5 papers used teachers as participants. However, researchers still also considered Twitter users and stakeholders for their study.

Table 6. The purposes of the studies.



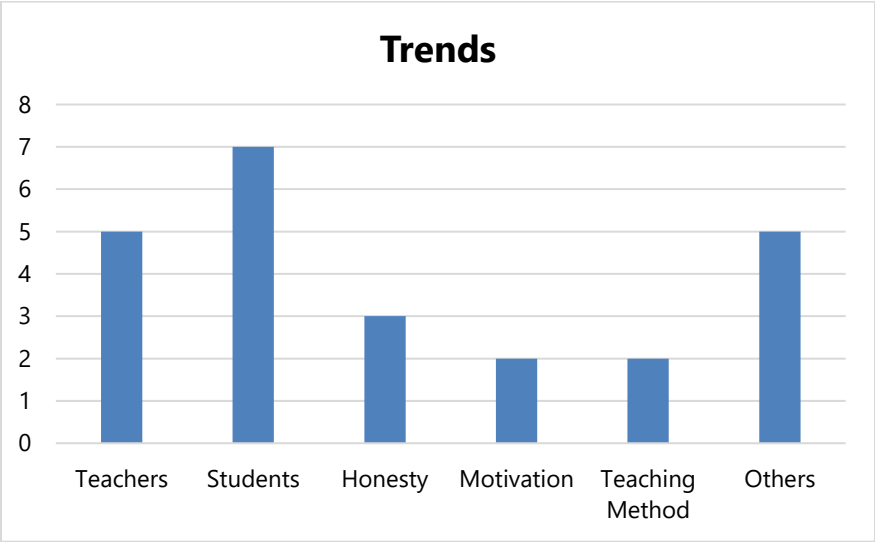
Ten papers have investigated or reviewed ChatGPT in education. Several other papers aimed to investigate how students or teachers utilize ChatGPT in education. Those two objectives are the most common in the reviewed papers.

Table 7. Problems in English language learning.

Problems	Number
Teachers perspective about ChatGPT in language learning	2
The honesty of students in language learning using ChatGPT	6
Digital competence of teachers regarding ChatGPT.	1
Concerns related to academic honesty and plagiarism.	2
Teachers' assessment of student exam results with the presence of ChatGPT	1
Student motivation and habits in accepting ChatGPT technology	2
Others	4
The potential to circumvent learning opportunities and contribute to the spread of misinformation	1
The quality of argumentative essays written by students using ChatGPT	1
Students perceived about ChatGPT	1
Challenge of achieving student-led learning in classroom setting that is primarily led by the teacher.	1
Teaching Method	2

The table above illustrates the issues in English language learning discussed in the reviewed papers. The most common problem encountered in language learning is the honesty of students when using ChatGPT. For instance, many students extensively utilize ChatGPT to complete assignments or essays and even to take exams with ChatGPT assistance. This poses a significant concern in the educational realm as students are perceived to lack critical thinking skills by relying solely on ChatGPT for task completion.

Table 8. Trends.



The "student" problem was most frequently observed (7 papers). For example, student honesty in education causes this issue to occur most frequently. Another problem that frequently appeared in the observed papers is "Teacher" (5 papers).

5. Conclusion

The meta-analysis study offers a substantial and valuable summary of publication patterns, publishing platforms, prolific authors, highly referenced articles, and sources within the ChatGPT domain. Moreover, it delves into the most frequently investigated topics concerning ChatGPT, analyzing their evolution over recent decades. This analysis aids in a deeper comprehension of pivotal ChatGPT-related issues, such as its performance and its application in teaching. Such research synthesis holds notable implications for task-oriented research, pedagogy, and methodological advancements.

In recent years, ChatGPT, one of the most controversial and widely used generative AI systems, has emerged as a potential game changer in the field of education, given its educational capabilities and excellence (Jeon & Lee, 2023; Kohnke et al., 2023; Tlili et al., 2023). ChatGPT has the potential to transform education by supporting various teaching methods and providing assistance to teachers. In the field of language, ChatGPT has made a great contribution, but its use in assessment creates ethical concerns about accuracy and bias. There are also concerns about ChatGPT perpetuating stereotypes and discrimination. Although it has not reached the expected level of performance, it is important to conduct further research to understand the impact as a whole, including student responses, ethical responses, and educational improvements.

The present study also has important implications for the potential of ChatGPT in the educational context. First, as artificial intelligence in the modern era, the topic in this study will surely inform and guide today's teaching and learning practices, which may, in turn, provide valuable feedback and impetus for further disciplinary development. Second, the topics that chatGPT in language learning is most interested in may overlap with issues that language educators are concerned with to some extent. For instance, the rising popularity of topics related to chatGPT in language learning partly reflects the effectiveness of AI in language teaching that educators can apply, besides how to use other digital media in learning. Third, the findings from this research synthesis also provide potential guidelines for language teaching, such as how accurate chatGPT is in writing an essay and the appropriate use of chatGPT in the classroom.

There are several methodological considerations for how studies like this should be conducted, especially in the context of meta-analysis. One possible methodological implication of this study is the use of meta-analysis to evaluate the performance of ChatGPT in various natural language tasks. Previous research has used methods such as topic analysis or topic modeling to evaluate the performance of ChatGPT, but these methods may face challenges in terms of manual assessment and interpretability. This study adopts meta-analysis and demonstrates that it is an effective and efficient approach to evaluating the performance of ChatGPT, with potential for further development. For example, future research could explore the use of additional evaluation metrics beyond accuracy and speed, as well as applying more advanced methods such as machine learning algorithms to enhance the proposed analysis.

This study only utilized abstracts of research articles to explore topics related to ChatGPT. Future studies may consider using full texts not only of research articles but also of other genres, such as monographs and book chapters, to provide a more comprehensive overview of research in this field. Additionally, we employed a list of ChatGPT-related search terms for data retrieval. Although this method helped us collect data more accurately and comprehensively, retrieval based on its application might overlook some emerging topics as the list of search terms may not be entirely exhaustive.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

[1] Ali, F., Choy, D., Divaharan, S., Tay, H. Y., & Chen, W. (2023). Supporting self-directed learning and self-assessment using TeacherGAIA, a generative AI chatbot application: Learning approaches and prompt engineering. Learning: *Research and Practice, 9*(2), 135-147.
[2] Barrot, J. S. (2023). ChatGPT as a language learning tool: An emerging technology report. Technology, Knowledge and Learning, 1-6.
[3] Božić, V., & Poola, I. (2023). Chat GPT and education. Preprint.
[4] Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. Innovations in Education and Teaching International, 61(2), 228-239.
[5] Chiu, T. K. (2023). The impact of Generative AI (GenAI) on practices, policies, and research direction in education: A case of ChatGPT and Midjourney. Interactive Learning Environments, 1-17.
[6] De Winter, J. C. (2023). Can ChatGPT pass high school exams on English Language Comprehension? *International Journal of Artificial Intelligence in Education,* 1-16.
[7] Divekar*, R. R., Drozdal*, J., Chabot*, S., Zhou, Y., Su, H., Chen, Y., ... & Braasch, J. (2022). Foreign language acquisition via artificial intelligence and extended reality: design and evaluation. *Computer Assisted Language Learning, 35*(9), 2332-2360.
[8] Elaish, M. M., Shuib, L., Ghani, N. A., & Yadegaridehkordi, E. (2019). Mobile English language learning (MELL): A literature review. *Educational Review, 71*(2), 257-276.
[9] Fütterer, T., Fischer, C., Alekseeva, A., Chen, X., Tate, T., Warschauer, M., & Gerjets, P. (2023). ChatGPT in education: global reactions to AI innovations. Scientific reports, 13(1), 15310.
[10] Farazouli, A., Cerratto-Pargman, T., Bolander-Laksov, K., & McGrath, C. (2023). Hello GPT! Goodbye home examination? An exploratory study of AI chatbots impact on university teachers' assessment practices. Assessment & Evaluation in Higher Education, 1-13.
[11] Fuchs, K. (2023, May). Exploring the opportunities and challenges of NLP models in higher education: is Chat GPT a blessing or a curse? In Frontiers in Education (8, 1166682). Frontiers.
[12] Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2023). A SWOT analysis of ChatGPT: Implications for educational practice and research. Innovations in Education and Teaching International, 1-15.
[13] Farhi, F., Jeljeli, R., Aburezeq, I., Dweikat, F. F., Al-shami, S. A., & Slamene, R. (2023). Analyzing the students' views, concerns, and perceived ethics about chat GPT usage. Computers and Education: Artificial Intelligence, 100180.
[14] Gorichanaz, T. (2023). Accused: How students respond to allegations of using ChatGPT on assessments. Learning: Research and Practice, 9(2), 183-196.
[15] Hopfenbeck, T. N. (2023). The future of educational assessment: Self-assessment, grit, and ChatGPT?. Assessment in Education: Principles, Policy & Practice, 30(2), 99-103.
[16] Herbold, S., Hautli-Janisz, A., Heuer, U., Kikteva, Z., & Trautsch, A. (2023). A large-scale comparison of human-written versus ChatGPT-generated essays. Scientific Reports, 13(1), 18617.
[17] Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. Relc Journal, 54(2), 537-550.
[18] Kalla, D., & Smith, N. (2023). Study and analysis of chat GPT and its impact on different fields of study. *International Journal of Innovative Science and Research Technology*, 8(3), 827-833.
[19] Mamo, Y., Crompton, H., Burke, D., & Nickel, C. (2024). Higher Education Faculty Perceptions of ChatGPT and the Influencing Factors: A Sentiment Analysis of X. TechTrends, 1-15.
[20] Newton, P., & Xiromeriti, M. (2023). ChatGPT performance on multiple choice question examinations in higher education. A pragmatic scoping review. Assessment & Evaluation in Higher Education, 1-18.
[21] Pokkakillath, S., & Suleri, J. (2023). ChatGPT and its impact on education. Research in Hospitality Management, 13(1), 31-34.
[22] Strzelecki, A. (2023). Students' acceptance of ChatGPT in higher education: An extended unified theory of acceptance and use of technology. Innovative Higher Education, 1-23.
[23] Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. Smart Learning Environments, 10(1), 15.
[24] Tan, S. C., Chen, W., & Chua, B. L. (2023). Leveraging generative artificial intelligence based on large language models for collaborative learning. Learning: Research and Practice, 9(2), 125-134.
[25] Ulla, M. B., Perales, W. F., & Busbus, S. O. (2023). 'To generate or stop generating response': Exploring EFL teachers' perspectives on ChatGPT in English language teaching in Thailand. Learning: Research and Practice, 9(2), 168-182.
[26] Wang, X., Pang, H., Wallace, M. P., Wang, Q., & Chen, W. (2022). Learners' perceived AI presences in AI-supported language learning: a study of AI as a humanized agent from community of inquiry. Computer Assisted Language Learning, 1-27.
[27] Xu, X., Dugdale, D. M., Wei, X., & Mi, W. (2023). Leveraging artificial intelligence to predict young learner online learning engagement. *American Journal of Distance Education*, 37(3), 185-198.
[28] Yu, H. (2023). Reflection on whether Chat GPT should be banned by academia from the perspective of education and teaching. Frontiers in Psychology, 14, 1181712.
[29] Zhao, X., Cox, A., & Cai, L. (2024). ChatGPT and the digitisation of writing. Humanities and Social Sciences Communications, 11(1), 1-9.
[30] Zhang, X., Sun, J., & Deng, Y. (2023). Design and Application of Intelligent Classroom for English Language and Literature Based on Artificial Intelligence Technology. Applied Artificial Intelligence, 37(1), 2216051.