Jurnal Pendidikan Bahasa Inggris Undiksha
Volume 12, Number 3, 2024 pp. 283-291
P-ISSN: 2614-1906 E-ISSN : 2614-1892
Open Access: https://ejournal.undiksha.ac.id/index.php/JPBI

Does ChatGPT Have a Significant Effect to Improve EFL Preservice Teachers' Teaching Plans? A Mixed-Method Study

Luh Gd Rahayu Budiarta[1*], I Putu Indra Kusuma[2]
[1,2] English Language Education, Universitas Pendidikan Ganesha, Singaraja, Indonesia

ARTICLE INFO

Article history:
Received October 05, 2024
Accepted December 11, 2024
Available online December 25, 2024

Kata Kunci:
ChatGPT, ChatGPT untuk mengajar, Guru Prajabatan EFL, Praktik Mengajar

Keywords:
ChatGPT, ChatGPT for Teaching, EFL Preservice Teacher, Teaching Practicums

DOI:
https://doi.org/10.23887/jpbi.v12i3.85769

ABSTRAK

Munculnya ChatGPT telah mengejutkan banyak pendidik bahasa Inggris, karena secara teoritis, ia memiliki banyak potensi untuk mendukung pengajaran bahasa Inggris. Namun, hasil empiris tentang bagaimana ChatGPT memengaruhi rencana pengajaran guru prajabatan bahasa Inggris sebagai bahasa asing (selanjutnya, EFL) masih belum jelas. Oleh karena itu, penelitian ini bertujuan untuk menyelidiki pengaruh ChatGPT terhadap rencana pengajaran guru prajabatan EFL. Penelitian ini menggunakan pendekatan metode campuran dan merekrut 17 guru prajabatan EFL untuk bergabung dalam penelitian. Data dikumpulkan menggunakan rubrik penilaian dan protokol wawancara. Dengan menggunakan Uji Mann-Whitney U dan analisis tematik induktif, penelitian ini menemukan bahwa rencana pelajaran yang dibuat oleh guru prajabatan EFL dalam kelompok eksperimen gagal mengungguli mereka dalam kelompok kontrol. Selain itu, rencana pelajaran pra tidak memiliki perbedaan yang signifikan dengan yang pasca. Data kualitatif menjelaskan hasil statistik. Kemudian, dua implikasi diambil untuk menggunakan ChatGPT secara efektif untuk mendukung pengajaran dan pembelajaran.

ABSTRACT

The advent of ChatGPT has surprised many English educators, as theoretically, it has many potentials to support English language teaching. However, the empirical results of how ChatGPT influence English as a foreign language (henceforth, EFL) preservice teachers' teaching plans remain unclear. This study therefore purposed to investigate the effect of ChatGPT towards EFL preservice teachers' teaching plans. This study employed a mixed-method approach and recruited 17 EFL preservice teachers to join the research. The data was collected using a scoring rubric and an interview protocol. Using *Mann-Whitney U Test* and inductive thematic analysis, this study found that the lesson plans created by EFL preservice teachers in the experimental group failed to outperform those in the control group. Moreover, the pre lesson plans had no significant difference with the post ones. The qualitative data explains the statistical results. Then, two implications were drawn to effectively use ChatGPT to support teaching and learning.

1. INTRODUCTION

Scholars have recognized that teaching practicums are a means to successfully practice teaching skills in a real-world setting (Altalhab et al., 2021; Opfer & Pedder, 2011; Safari, 2020). Preservice teachers use the opportunity to develop their identity as educators through teaching practicums as they would implement the knowledge they attained from their teacher education program (Batane & Ngwako, 2017; Safari, 2020). Thus, teaching practicums are necessary for preservice teachers' identity development. Previous studies have indicated that English as a foreign language (henceforth, EFL) preservice teachers faced some obstacles when doing teaching practicums (Azimi et al., 2019; Batane & Ngwako, 2017; Baz et al., 2018; Mudra, 2018). For example, Indonesian EFL preservice teachers had issues with finding appropriate teaching media and materials (Chasanah & Sumardi, 2022; Merç, 2015; Mudra, 2018). Despite the fact that EFL preservice teachers frequently encountered obstacles due to their lack of knowledge and teaching experience, current EFL teachers who are acquittanced with the use of technology frequently seek for solutions on some technology platforms, such as Google and YouTube (Kusuma, 2023; Opfer & Pedder, 2011; Salinas & Ayala, 2018). In addition, the advent of current technological tools, such as generative artificial intelligence (henceforth, AI) applications such as ChatGPT that can generate information like Google, have the potential to provide solutions about teaching and learning issues. ChatGPT, an AI-powered chatbot, was developed by OpenAI as a natural language processing system (Kalla & Smith, 2023; Nghi & Phuc, 2023). ChatGPT can perform various tasks and generate information required by the users (Cheong & Hong, 2023;

** Corresponding Author: Luh Gd Rahayu Budiarta: rahayu.budiarta@undiksha.ac.id*

Cotton et al., 2023; Fitria, 2023). Consequently, ChatGPT can be a useful resource and platform for teachers, including preservice ones, to search for information that will assist them in becoming more professional teachers (Farrokhnia et al., 2023; Kohnke et al., 2023).

The quality of schools and education will only improve if the quality of its educators improves as well (Opfer & Pedder, 2011; Spyropoulou & Kameas, 2024). Teachers already in the classroom would be able to further their education and professional development through the lessons they teach and the workshops they attend. Preservice teachers, on the other hand, get improved due to their coursework and teaching practicums. As preservice teachers would implement what they have learned from teacher education program, teaching practicums are a means to successfully practice teaching skills in a real-world setting (Altalhab et al., 2021; Batane & Ngwako, 2017). Teaching practicums also provide opportunities for preservice teachers to grow their identity as teachers. Thus, teaching practicums are necessary for preservice teachers' development (Kwaah et al., 2022; Safari, 2020). Previous research has shown that preservice teachers, the majority of whom are inexperienced and untrained in the field of education, frequently run into difficulties when completing their teaching practicums. For example, preservice teachers found differences on what they learned in teacher education program and what they found in the real schools (Salazar Noguera & McCluskey, 2017), difficulties when designing assessment (Azimi et al., 2019), issues to find and create appropriate teaching media and materials, issues using English as a medium of instruction (Chasanah & Sumardi, 2022), issues with technology integration to support English instruction, and problems with teaching management. The transition from being students to teachers is probably the source of the problems. As a consequence, preservice teachers would experience low self-confidence during teaching (Nugroho, 2017; Permatasari et al., 2019).

Technological advances have an impact on all fields, one of which has an impact on the field of language education. It is undeniable that technology has always been a part of English language education (N. J. Kim & Kim, 2022; Yue et al., 2022). Then, as technology develops, the arrival of artificial intelligence (AI) also brings renewal in education, especially in language learning (Gonulal, 2021; Huang et al., 2023; Liang et al., 2021; Tai & Chen, 2022; Tsai, 2019). For example, the use of Google Translate to help students translate their native language to the target one and the use of artificial intelligence personal assistance like Google Home Hub and Home Mini could help the students learn and develop their speaking skills (Kurniawati et al., 2019; Turgut, 2017). As a result, the advent of AI has resulted in the introduction of important reforms in language learning. ChatGPT, a sophisticated AI-powered language model, has attracted considerable attention in a remarkably brief period of time since its launch (Cheong & Hong, 2023; Sallam, 2023). This AI system can translate from a user's native language to a desired target language, generate written content in the target language, and engage in interactive conversations with users. As a result, these remarkable characteristics of ChatGPT have the potential to cause significant disruptions in numerous industries, including English language instruction (Afkarin & Asmara, 2024; Fitria, 2023; Shikun et al., 2024). The emergence of ChatGPT, with its extraordinary capacity to generate natural and coherent texts (Nghi & Phuc, 2023), has posed new challenges in the field of education as many educators do not know how to respond to this new AI technology, resulting in extensive discussions and debates (Chaudhry et al., 2023; Lo, 2023). In addition, ChatGPT has raised concerns about its potential impact on the development of students' writing abilities. Concerns exist that students may cheat on online examinations, generate texts that closely resemble human language, experience a decline in critical thinking skills, and struggle to evaluate information (Kjell et al., 2024; Rahman & Watanobe, 2023). Moreover, previous study found that the texts generated by ChatGPT could not be fully detected by AI detector applications (Constantia et al., 2021). These perceived risks therefore contribute to an increase in educators' and students' reluctance to use ChatGPT for English language instruction.

In contrast, research indicates that ChatGPT has the potential to substantially improve language instruction. ChatGPT can generate a complete response when a user enters a query or a sentence fragment. Then, text entered by the users in their native language is translated automatically into the target language. ChatGPT is smart enough to respond to follow-up inquiries, facilitating personalized learning and complex learning (Guzik et al., 2023; Kjell et al., 2024; Shikun et al., 2024). In addition, ChatGPT can churn out humor and poetry, among other kinds of written expression. In addition, users can access dictionary definitions and examples via ChatGPT. For instance, it can define a term, identify its grammatical category, provide example sentences, and provide additional definitions. ChatGPT, is also known for its ability to write English essays with appropriate grammar patterns (Cotton et al., 2023; Farrokhnia et al., 2023; Sallam, 2023). Regarding its potential for teachers, ChatGPT can also be implemented for creating teaching materials and help the teachers with scoring tasks, decreasing teachers' workload (Koraishi, 2023; Lo, 2023). However, with the limited information provided by previous research, the information about whether ChatGPT can help improve EFL teachers' instructions, particularly the preservice teachers who lack teaching experiences. This information is necessary since ChatGPT has great potentials for education, especially as a teaching assistant for EFL preservice teachers. It is therefore pivotal to conduct this research to find out whether ChatGPT can help improve EFL preservice teachers' teaching plans or

not. Since ChatGPT is new to education, such information will bring some novelties for English education. The aims of this study is to investigate the effect of ChatGPT towards EFL preservice teachers' teaching plans.

2. METHOD

The researchers used a sequential explanatory mixed method in which quantitative research was carried out first and then followed by qualitative research. This research employed *pre-experimental* with *static group comparison* design for the quantitative part. Pre-experimental was employed because the inability of recruiting enough for doing bigger experimental study, such as those in quasi experimental or true experimental study. In the *static group comparison*, the researchers only provided treatments to experimental participants and did not do anything to those in control group (Gopalan et al., 2020). Moreover, the qualitative research used in this study is a basic qualitative approach, which is to conduct qualitative research in qualitative ways without having to worry about what approach is actually suitable for this qualitative research. The study took place in Indonesia, particularly at a state university that has English Language Education Program. The context of this research was the effect of ChatGPT towards Indonesian EFL preservice teachers' teaching plans.

The participants recruited were EFL preservice teachers in the English Education Study Program. After obtaining research permission, the researcher contacted the participants to explain the study. Potential participants were recruited using convenience sampling technique which looked for participants who were available and willing to spend time (Ary et al., 2019; Herbein et al., 2018). The researcher recruited 20 students to form 2 groups, namely, a group that received treatment and a group without receiving treatment. The researchers, before recruiting the study, asked permission from the coordinator of the English Language Education study program to ask permission to approach the EFL preservice teachers who were doing teaching practice. Of all EFL preservice teachers who had been approached, initially 20 of them were found to be willing and then they were divided into 2 groups with ten EFL preservice teachers in each. Then, halfway through, 3 of them in the experimental group withdrew due to fear of not being able to participate optimally. Therefore, only seven EFL preservice teachers followed until the end in the experimental class and ten in the control class. Then, after the experiment was completed, these seven participants were invited to conduct interviews to explore the data obtained from the experimental data.

The Researchers collected quantitative data through tests by collecting lesson plans or teaching modules before and after treatment which are assessed using an assessment rubric which had been validated by two experts in English language teaching. More specifically, during the initial week of the study, the participants were asked to submit their initial lesson plans, which were prepared without the help of ChatGPT. After the end of the study, and before the end of the teaching practice period, the participants were again assigned to submit their final lesson plans after receiving training using ChatGPT. Next, the researcher conducted a comparative analysis of the two sets of lesson plans to see important differences in their content and structure. We then conducted a non-parametric comparison test, specifically the *Mann-Whitney-U test*, to determine the results of the two lesson plans. After obtaining the results of the comparative test, the researchers determined how the participants' ability to plan lessons improved or not. Then the researchers conducted semi-structured interviews guided by the interview protocol which had been validated. The interview process included some research-related questions about the impact of ChatGPT on their teaching career and how ChatGPT supported their professionalism. The researchers utilized the interviews to find out more about how the prospective English teachers' skills improved after using ChatGPT to develop their pedagogical innovations, such as classroom management skills, lesson design skills, and so on. To reduce fear and increase the possibility of getting more thorough information, the researcher asked questions based on whether the participants felt comfortable speaking in Indonesian or English, and the interviews would be tailored to the EFL teachers' needs for fluent dialog.

To analyze the quantitative data, the researcher applied non-parametric statistics such as *dispersion* and non-parametric test using the *Mann-Whitney U Test*. The *Mann-Whitney U Test* was used in this study because the number of participants included in a class was very limited and it was not possible to conduct prerequisite tests such as normality and homogeneity. The researchers employed SPSS to help analyze the data. Regarding the qualitative data, the interview data will be transcribed first. Then the data will be read and coded using in-vivo techniques and continued with inductive thematic analysis.

3. RESULT AND DISCUSSION

Result

The first hypothesis test determines whether there is a significant difference between experimental participants' post lesson plan (ChatGPT-assisted) and control participants' post lesson plan during teaching practice in terms of several variables, such as learning materials, learning strategies and methods, learning ideas, learning

activities, and learning assessment. Lesson plan assessment results between experimental and control lesson plans is show in Table 1.

Table 1. Lesson Plan Assessment Results Between Experimental and Control Lesson Plans

Lesson Plans	N	Mean Rank	Sum of Ranks
Post (Exp)	7	5.21	36.50
Post (Control)	1	11.65	116.50
Total	1		

The statistical findings in Table 1 show that the control participants' post lesson plan (*N*=10) has a *mean rank* = 11.65 and *sum of ranks* = 116.50 which is slightly better than the experimental participants' post lesson plan (*N*=7) which has a *mean rank* = 5.21 and *sum of ranks* = 36.50. However, the *Mann-Whitney U* results is show in Table 2.

Table 2. Participants' experimental post-learning plans (assisted by chatGPT) with control post-learning plans

Mann Whitney U	Wilcoxon W	Z	Asymp. Sig. (2-tailed)
8.500	36.500	-2.633	0.008

Base on Table 2 show that the post-experimental lesson plan and the post-control lesson plan have significant differences in the pre-assessed lesson plan section, indicating that the post-control lesson plan outperforms the post-experimental lesson plan with a *Mann Whitney-U* value = 8.500 and *Asymp. Sig (2-tailed)* = $p < 0.008$. In addition, researchers conducted a specific comparison test using the *Mann-Whitney U test* to obtain more detailed data covering all learning implementation plan variables which can be seen in Table 3.

Table 3. Special Comparison Test Results (Experimental and Control Lesson Plans)

Variables	Mann Whitney U	Wilcoxon W	Z	Asymp.Sig.(2-tailed)
Learning objectives	10.000	38.000	-3.056	0.002
Strategy & teaching methods	25.000	53.000	-1.317	0.188
Assessments	4.000	32.000	-3.175	0.001
Learning materials	27.500	55.500	-857	0.391
Teaching procedure	8.000	36.000	-3.033	0.002

Base on Table 3, the data shows that the lessons plan from the experimental and control groups had significant different on learning objectives, assessments, and teaching procedure where according to Table 4 the scores in the control group outperformed those in the experimental group.

Table 4. The Results Score of the Lessons Plan from the Experimental and Control Groups

Variables	Lesson Plan	Mean	Std. Deviation	Mean Rank	N
Learning objectives	Post (exp)	13.1429	1.95180	5.43	7
	Post (control)	16.0000	0.00000	11.50	10
	Total	14.8235	1.87867		17
Strategy & teaching methods	Post (exp)	10.8571	3.02372	7.57	7
	Post (control	12.0000	0.00000	10.00	10
	Total	11.5294	1.94029		17
Assessments	Post (exp)	10.2857	4.53557	4.57	7
	Post (control)	18.4000	2.06559	12.10	10
	Total	15.0588	5.20181		17
Learning materials	Post (exp)	13.7143	3.14718	7.93	7
	Post (control)	14.0000	2.10819	9.75	10
	Total	13.8824	2.49706		17

Variables	Lesson Plan	Mean	Std. Deviation	Mean Rank	N
Teaching procedure	Post (exp)	16.0000	2.30940	5.14	7
	Post (control)	19.6000	1.26491	11.70	10
	Total	18.1176	2.49706		17

The researcher continued by conducting comparative tests on the participants' pre- and post-experiment lesson plans (assisted by ChatGPT) to determine if there was a significant difference between the two lesson plans. The results showed that there was no significant difference between the participants' experimental lesson plans and the post-experimental lesson plans (using ChatGPT). Statistical results between participants' experimental pre lesson plans and post lesson plans is show in Table 5.

Table 5. Statistical results between participants' experimental pre lesson plans and post lesson plans (using ChatGPT)

Lesson Plans	N	Mean Rank	Sum of Ranks
Pre (Exp)	7	8.64	60.50
Post (Exp)	7	6.36	44.50
Total	14		

The statistical findings in Table 5 show that the participants' experimental lesson plans (N = 7) had a *mean rank* = 8.64 and *sum of ranks* = 60.50 which was slightly better than the post-experimental lesson plans (N = 7), which had a *mean rank* = 6.36 and *sum of ranks* = 44.50. The U test result is show in Table 6.

Table 6. Mann-Whitney U test' Results

Mann Whitney U	Wilcoxon W	Z	Asymp. Sig. (2-tailed)
16.500	44.500	-1.049	0.294

Base on Table 6 show that participants' pre- and post-experimental lesson plans did not have any significant differences in the pre-assessed portion of the lesson plan, indicating that participants' pre-experimental lesson plans were quite similar to participants' post-experimental lesson plans (those using ChatGPT) with a *Mann Whitney-U* value = 16.500 and *Asymp. Sig* (2-tailed) = $p > 0.294$. Then, the result of qualitative result is show in Table 7.

Table 7. Themes and Sub-themes

Theme	Sub-themes	Partisipants (n)
EFL preservice teachers are too indifferent to lesson plans made by ChatGPTs	Limited time	1
	Overconfidence in ChatGPT capabilities	2
	Lack of Understanding of the Review Process	3
	Limited Information from ChatGPT.	1

The qualitative data in the Table 7 shows that the statements made by each participant regarding the assessment results of their final lesson plans, specifically what caused their first lesson plans to be better than their final lesson plans, created with the help of ChatGPT. The cause was a combination of various issues, including lack of time, overconfidence in ChatGPT's capabilities, lack of understanding of the review process, and limited information from ChatGPT. Proper control of ChatGPT is essential for creating effective lesson plans, as explained earlier in the answer to the first research question. So, no matter how good the idea is, it should be checked and adjusted to the needs of the students in the classroom in order to be implemented properly.

One participant said that they were under time constraints or other constraints during the teaching process, which limited their ability to complete a full examination of the ChatGPT lesson plans. The following is her explanation:

> *"I also thought that the lesson plans produced by ChatGPT did not need to be reviewed first because it would be used for comparison, then at that time I also had final assignments to do, so I did not have time to review it". (P3)*

Two participants were confident in the ChatGPT's ability to produce adequate lesson plans. As a result, they may not have felt the need to conduct extensive evaluation as they already felt that the results from ChatGPT were appropriate. The following is their explanation:

> *"In my opinion, I totally agree with you just now, when making lesson plans from ChatGPT, I just copy and paste and don't review at all, so the lesson plans that I have made without using ChatGPT are better than the final lesson plans"*. (P1)

Three participants may have underestimated the importance of conducting a thorough evaluation to ensure that the lesson plans are appropriate for the students' needs and classroom conditions. Some participants stated that they had not reviewed the lesson plans created by ChatGPT. Time constraints and other tasks, such as completing the KM program, made them hesitant or unable to conduct a thorough evaluation of the lesson plans. One participant said that the lesson plans created by ChatGPT provided general and not very specific information about the needs of students, schools and learning environments. Here is her explanation:

> *"The lesson plans generated by ChatGPT show an answer that is lacking because of course ChatGPT does not know who and where the lesson plans we make will be used for. The lesson plans generated by ChatGPT are only general explanations, so we have to adjust to the circumstances around us."* (P5)

Discussion

Testing the first hypothesis showed interesting dynamics when comparing experimental and control post-learning plans. Although the data showed a slight edge in average and total ratings for the control plans, the *Mann-Whitney U test* found a significant difference. This difference calls into question ChatGPT's ability to generate lesson plans when compared to teacher-generated lesson plans. This results therefore challenge the claims made previous researchers who said that ChatGPT had the potentials to improve the teachers' teaching plans (Kohnke et al., 2023; Sallam, 2023). Moreover, the specific results showed significant differences between the experimental and control groups in learning objectives, assessments, and teaching procedure which are vital parts of teaching plans. We surmised that not all information provided by the ChatGPT is relevant as claimed by previous researchers that it has the potential to generate biased information because it uses a large database that has not been confirmed its validity (Cheong & Hong, 2023; Sallam, 2023).

We also predicted that ChatGPT did not know what the school situation was like, what the students' needs were, what the school facilities were, or the characteristics of each student. So, if the lesson plans created by ChatGPT are not reviewed or modified to suit the needs of the class, the results will be disappointing as they do not match the needs of the students or the conditions of the school. In this case, it is important to check and modify the lesson plans created by ChatGPT to meet the needs of students and schools.

However, another concern also appears, where we also predicted that the participants also played important roles during the treatments. According to the interviews with the participants, they confessed that they did not review the lesson plans produced by ChatGPT, but rather just copied and pasted and used them directly without reviewing and adjusting to the needs of the students and the school situation (Guzik et al., 2023; Kjell et al., 2024). Time constraints due to other activities, overconfidence in ChatGPT's capabilities, lack of understanding in evaluating lesson plans, and ChatGPT's limitations in providing information, were the participants' reasons why their post-experiment lesson plans were not better than their pre-experiment lesson plans (Kalla & Smith, 2023; Sallam, 2023). Thus, even though scholars claimed that ChatGPT had potential to help teachers, without participants serious assessments, the information provided by ChatGPT will be useless.

The findings from this study carry significant implications for the integration of AI tools, specifically ChatGPT, into the teaching plans process. While ChatGPT and other AI-powered tools have been widely praised for their potential to assist teachers, this study challenges these claims, revealing critical gaps in the actual performance of AI-generated lesson plans. Firstly, AI as a Supplement, not a replacement for teacher expertise. The study highlights that AI tools like ChatGPT may not yet be ready to fully replace teacher-generated lesson plans, especially in areas such as learning objectives, assessments, and teaching procedures. Teachers bring contextual knowledge about students, school conditions, and curriculum standards that AI lacks. Therefore, AI should be viewed as a supplemental tool that requires the teacher's review and expertise to tailor lesson plans to the specific needs of students and the learning environment (Cotton et al., 2023; Farrokhnia et al., 2023). Second, A crucial implication is the necessity for teachers to thoroughly review and modify AI-generated lesson plans to align with their classroom realities. As the participants admitted, the failure to adapt and customize the AI-generated plans led to poorer outcomes. This underscores the importance of human intervention in ensuring that AI-generated content is relevant and meets the instructional goals of the classroom. Teachers must be equipped with the skills to critically assess and refine AI outputs, rather than relying on them unquestioningly because AI could provide generic information or even fake information (E.-J. Kim et al., 2021; N. J. Kim & Kim, 2022).

Despite the insightful findings from this study, several limitations were acknowledged. First, one significant limitation of this study is the small number of participants. With only a few participants, the findings may not be

generalizable to broader populations of EFL preservice teachers. Future studies should aim to involve a larger and more diverse sample of participants from different regions or institutions to strengthen the generalizability of the results and provide a more comprehensive understanding of AI's impact on teaching plans. Second, this study employed a pre-experimental design, which limits the ability to make causal inferences between the use of AI (ChatGPT) and the observed outcomes in teaching plans. Future research should consider employing a more rigorous research design.

4. CONCLUSION

This study found that the lesson plans created by EFL preservice teachers in the experimental group failed to outperform those in the control group. Moreover, the pre lesson plans had no significant difference with the post ones. The qualitative data explains the statistical results. While ChatGPT offers promising tools for educational innovation, their successful use in teaching relies heavily on teacher expertise, critical review, and contextual adaptation. Without these, the use of AI in teaching plans risks falling short of its potential and may even lead to poorer educational outcomes. This study calls for a more nuanced approach to integrating AI in education, one that emphasizes the indispensable role of teachers as informed facilitators and critical evaluators of AI-generated content.

5. REFERENCES

Afkarin, M. Y., & Asmara, C. H. (2024). Investigating the Implementation of ChatGPT in English Language Education: Effects on Student Motivation and Performance Levels. *Journey: Journal of English Language and Pedagogy*, *7*(1), 57–66. http://ejurnal.uibu.ac.id/index.php/journey/article/view/865.

Altalhab, S., Alsuhaibani, Y., & Gillies, D. (2021). The reflective diary experiences of EFL pre-service teachers. *Reflective Practice*, *22*(2), 173–186. https://doi.org/10.1080/14623943.2020.1865903.

Ary, D., Jacobs, L. C., Sorensen, C. K., & Walker, D. A. (2019). *Introduction to research in education* (10th ed.). Cengage.

Azimi, E., Kuusisto, E., Tirri, K., & Hatami, J. (2019). How do student teachers reflect on their practice through practicum courses? A case study from Iran. *Journal of Education for Teaching*, *45*(3), 277–289. https://doi.org/10.1080/09589236.2019.1599511.

Batane, T., & Ngwako, A. (2017). Technology use by pre-service teachers during teaching practice: Are new teachers embracing technology right away in their first teaching experience? *Australasian Journal of Educational Technology*, *33*(1), 48–61. https://doi.org/10.14742/ajet.2299.

Baz, H. E., Balçıkanlı, C., & Cephe, P. T. (2018). Introducing an innovative technology integration model: Echoes from EFL pre-service teachers. *Education and Information Technologies*, *23*(5), 2179–2200. https://doi.org/10.1007/s10639-018-9711-9.

Chasanah, I., & Sumardi, S. (2022). A retrospective narrative study of pre-service teachers' language barrier experiences in international teaching practicum. *VELES Voices of English Language Education Society*, *6*(1), 271–282. https://doi.org/10.29408/veles.v6i1.5356.

Chaudhry, I. S., Sarwary, S. A. M., El Refae, G. A., & Chabchoub, H. (2023). Time to revisit existing student's performance evaluation approach in higher education sector in a new era of ChatGPT: A case study. *Cogent Education*, *10*(1), 1–30. https://doi.org/10.1080/2331186X.2023.2210461.

Cheong, W., & Hong, H. (2023). The impact of ChatGPT on foreign language teaching and learning : Opportunities in education and research. *Journal of Educational Technology and Innovation*, *3*(1), 37–45. https://doi.org/10.61414/jeti.v5i1.103.

Constantia, C., Christos, P., Glykeria, R., Anastasia, A., & Aikaterini, V. (2021). The Impact of COVID-19 on the Educational Process : The Role of the School Principal. *Journal of Education*, 1–8. https://doi.org/10.1177/00220574211032588.

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 1–12. https://doi.org/10.1080/14703297.2023.2190148.

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*. https://doi.org/10.1080/14703297.2023.2195846.

Fitria, T. N. (2023). Artificial intelligence (AI) technology in OpenAI ChatGPT application: A review of ChatGPT in writing English essay. *Journal of English Language Teaching*, *6*(1), 44–58. https://journal.unnes.ac.id/sju/elt/article/view/64069.

Gonulal, T. (2021). Investigating EFL learners' humorous interactions with an intelligent personal assistant.

Interactive Learning Environments. https://doi.org/10.1080/10494820.2021.19744.

Gopalan, M., Rosinger, K., & Ahn, J. Bin. (2020). Use of Quasi-Experimental Research Designs in Education Research: Growth, Promise, and Challenges. *Review of Research in Education*, *44*(1), 218–243. https://doi.org/10.3102/0091732X20903302.

Guzik, E. E., Byrge, C., & Gilde, C. (2023). The originality of machines: AI takes the Torrance Test. *Journal of Creativity*. https://doi.org/10.1016/j.yjoc.2023.100065.

Herbein, E., Golle, J., Tibus, M., Schiefer, J., Trautwein, U., & Zettler, I. (2018). Fostering elementary school children's public speaking skills: A randomized controlled trial. *Learning and Instruction*, *55*(October 2017), 158–168. https://doi.org/10.1016/j.learninstruc.2017.10.008.

Huang, X., Zou, D., Cheng, G., Chen, X., & Xie, H. (2023). Trends, research issues and applications of artificial intelligence in language education. *Technology & Society*, *26*(1), 112–131. https://doi.org/10.2307/48707971.

Kalla, D., & Smith, N. B. (2023). Study and analysis of ChatGPT and its impact on different fields of study. *International Journal of Innovative Science and Research Technology*, *8*(3), 827–833. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4402499.

Kim, E.-J., Kim, J. J., & Han, S.-H. (2021). Understanding Student Acceptance of Online Learning Systems in Higher Education: Application of Social Psychology Theories with Consideration of User Innovativeness. In *Sustainability* (Vol. 13, Issue 2). https://doi.org/10.3390/su13020896.

Kim, N. J., & Kim, M. K. (2022). Teacher's perceptions of using an artificial intelligence-based educational tool for scientific writing. *Frontiers in Education*, *7*, 1–13. https://doi.org/10.3389/feduc.2022.755914.

Kjell, O. N. E., Kjell, K., & Schwartz, H. A. (2024). Beyond rating scales: With targeted evaluation, large language models are poised for psychological assessment. In *Psychiatry Research* (Vol. 333). https://doi.org/10.1016/j.psychres.2023.115667.

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, 1–14. https://doi.org/10.1177/00336882231162868.

Koraishi, O. (2023). Teaching English in the age of AI: Embracing ChatGPT to optimize EFL materials and assessment. *Language Education & Technology (LET Journal)*, *3*(1), 55–72. https://langedutech.com/letjournal/index.php/let/article/view/48.

Kurniawati, M., Santanapurba, H., & Kusumawati, E. (2019). Penerapan Blended Learning Menggunakan Model Flipped Classroom Berbantuan Google Classroom Dalam Pembelajaran Matematika Smp. *EDU-MAT: Jurnal Pendidikan Matematika*, *7*(1), 8–19. https://doi.org/10.20527/edumat.v7i1.6827.

Kusuma, I. P. I. (2023). How do Indonesian EFL student teachers solve teaching issues during online teaching practicums? A phenomenological study. In *Handbook of CALL Teacher Education and Professional Development* (pp. 99–115). Springer Nature Singapore. https://doi.org/10.1007/978-981-99-0514-0_7.

Kwaah, C. Y., Adu-Yeboah, C., Amuah, E., Essilfie, G., & Somuah, B. A. (2022). Exploring preservice teachers' digital skills, stress, and coping strategies during online lessons amid covid-19 pandemic in Ghana. *Cogent Education*, *9*(1), 2107292. https://doi.org/10.1080/2331186X.2022.2107292.

Liang, J.-C., Hwang, G.-J., & Chen, M.-R. A. (2021). Roles and research foci of artificial intelligence in language education: An integrated bibliographic analysis and systematic review approach. *Interactive Learning Environments*. https://doi.org/10.1080/10494820.2021.19583.

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, *13*(410), 1–15. https://doi.org/10.3390/educsci13040410.

Merç, A. (2015). The Potential of General Classroom Observation: Turkish EFL Teachers' Perceptions, Sentiments, and Readiness for Action. *Journal of Education and Training Studies*, *3*(4). https://doi.org/10.11114/jets.v3i4.821.

Mudra, H. (2018). Pre-service EFL teachers' experiences in teaching practicum in rural schools in Indonesia. *Qualitative Report*, *23*(2), 319–344. https://doi.org/10.46743/2160-3715/2018.3115.

Nghi, T. T., & Phuc, T. huu. (2023). Bots as foreign language learning tools: From theory to practice. *Revista De Educacion (Madrid)*, *389*(9), 13–21. https://www.researchgate.net/profile/Tran-Nghi-4/publication/344378131_Bots_as_Foreign_Language_Learning_Tools_From_Theory_To_Practice/links/5f6dee61458515b7cf4c7149/Bots-as-Foreign-Language-Learning-Tools-From-Theory-To-Practice.pdf.

Nugroho, H. A. (2017). Preservice EFL teachers' self-efficacy, their English proficiency and their preparedness for teaching practicum program. *Premise Journal*, *6*(2), 1–11. https://ojs.fkip.ummetro.ac.id/index.php/english/article/view/997.

Opfer, V. D., & Pedder, D. (2011). Conceptualizing teacher professional learning. *Review of Educational Research*, *81*(3), 376–407. https://doi.org/10.3102/0034654311413609.

Permatasari, D., Mulyono, H., & Ferawati. (2019). Investigating the contributing factors to teaching anxiety during teaching practicum: A case of Indonesian pre-service EFL teachers. *Indonesian Research Journal in Education*, *3*(2), 236–248. https://online-journal.unja.ac.id/irje/article/view/7422.

Rahman, M. M., & Watanobe, Y. (2023). ChatGPT for education and research: Opportunities, threats, and strategies. *Applied Sciences*, 1–21. https://doi.org/10.3390/13095783.

Safari, P. (2020). Iranian ELT student teachers' portrayal of their identities as an English language teacher: Drawings speak louder than words. *Journal of Language, Identity and Education*, *19*(2), 125–141. https://doi.org/10.1080/15348458.2019.1650279.

Salazar Noguera, J., & McCluskey, K. (2017). A case study of early career secondary teachers' perceptions of their preparedness for teaching: lessons from Australia and Spain. *Teacher Development*, *21*(1), 101–117. https://doi.org/10.1080/13664530.2016.1204353.

Salinas, D., & Ayala, M. (2018). EFL Student-Teachers' Identity Construction: A Case Study in Chile. *HOW*, *25*(1), 33–49. https://doi.org/10.19183/how.25.1.380.

Sallam, M. (2023). ChatGPT utility in healthcare education, research, and practice: Systematic review on the promising perspectives and valid concerns. *Healthcare (Switzerland)*, *11*(6), 1–20. https://doi.org/10.3390/healthcare11060887.

Shikun, S., Grigoryan, G., Huichun, N., & Harutyunyan, H. (2024). AI Chatbots: Developing English Language Proficiency in EFL Classroom. *Arab World English Journal (AWEJ) Special Issue on ChatGPT*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4814753.

Spyropoulou, N., & Kameas, A. (2024). Augmenting the Impact of STEAM Education by Developing a Competence Framework for STEAM Educators for Effective Teaching and Learning. *Education Sciences*, *14*(1). https://doi.org/10.3390/educsci14010025.

Tai, T. Y., & Chen, H. H. J. (2022). The impact of intelligent personal assistants on adolescent EFL learners' speaking proficiency. *Computer Assisted Language Learning*, 1–28. https://doi.org/10.1080/09588221.2022.2070219.

Tsai, S. C. (2019). Using google translate in EFL drafts: a preliminary investigation. *Computer Assisted Language Learning*, *32*(5–6), 510–526. https://doi.org/10.1080/09588221.2018.1527361.

Turgut, Y. (2017). Tracing preservice English language teachers' perceived TPACK in sophomore, junior, and senior levels. *Cogent Education*, *4*(1). https://doi.org/10.1080/2331186X.2017.1368612.

Yue, M., Jong, M. S. Y., & Dai, Y. (2022). Pedagogical Design of K-12 Artificial Intelligence Education: A Systematic Review. *Sustainability (Switzerland)*, *14*(23), 1–29. https://doi.org/10.3390/su142315620.