ISSN 1799-2591
Theory and Practice in Language Studies, Vol. 14, No. 12, pp. 3851-3860, December 2024
DOI: https://doi.org/10.17507/tpls.1412.20

From Nervous to Fluent: The Impact of AI Chatbot-Assisted Assessment on English Reading Anxiety and Performance in Indonesia

Wiyaka Wiyaka*
English Education Department, Universitas PGRI Semarang, Semarang, Indonesia

Lusia Maryani Silitonga
English Education Department, Universitas PGRI Semarang, Semarang, Indonesia

Sunardi Sunardi
English Department, Universitas Dian Nuswantoro, Semarang, Indonesia

Yuventius Tyas Catur Pramudi
Information System Department, Universitas Dian Nuswantoro, Semarang, Indonesia

***Abstract*—This study examines the influence of AI Chatbot-assisted assessment on English reading anxiety and performance in Indonesian secondary school pupils. The growing integration of artificial intelligence in educational settings, namely through reading bots, has the potential to improve language acquisition and reduce student anxiety. The study utilizes a quantitative research approach, which compares the performance of an experimental group that interacts with AI chatbot-assisted assessment to a control group that receives traditional teaching led by a teacher. The data were gathered by administering pre- and post-tests of the adopted EFLRAI to assess reading anxiety, and the KET to measure reading performance. The results indicate that engagement with AI chatbots significantly alleviates reading anxiety and improves reading performance. This fosters a conducive learning atmosphere that enhances confidence and engagement in acquiring the English language. This study improves the understanding of the role of AI Chatbot-assisted assessment in language education and its effectiveness in addressing the challenges faced by English as a Foreign Language (EFL) learners in Indonesia.**

***Index Terms*—AI Chatbot, reading anxiety, reading performance, assessment, English learning**

I. Introduction

The rapid advancement and proliferation of artificial intelligence (AI) technologies have precipitated a paradigm shift across various domains, with the field of education experiencing particularly transformative effects (Holmes et al., 2019). Within this broader context, the application of AI in language pedagogy has emerged as a subject of considerable scholarly interest and practical exploration. Integrating AI-driven tools in language learning environments can revolutionize traditional instructional methodologies, offering personalized, adaptive, and interactive learning experiences that were hitherto unfeasible (Chapelle & Sauro, 2017).

Among the myriad AI applications in educational technology, conversational agents, colloquially known as chatbots, have garnered significant attention from researchers and practitioners alike. These sophisticated software programs present a promising avenue for enhancing language acquisition processes and mitigating learner anxiety (Fryer & Carpenter, 2006). The capacity of chatbots to provide immediate, contextually relevant feedback, engage in goal-oriented dialogues, and offer unlimited opportunities for language practice renders them particularly well-suited to addressing the multifaceted challenges inherent in second language acquisition (Bibauw et al., 2019).

In English learning, AI chatbots are essential as they offer tailored, interactive, and accessible language practice for learners. These chatbots can replicate authentic discussions, allowing students to hone their language abilities in a low-stress setting, so alleviating anxiety and fostering confidence. AI chatbots provide immediate feedback and personalized learning trajectories, addressing specific learning challenges while enabling students to advance at their own speed. Incorporating AI in English teaching improves the learning experience, rendering it more dynamic, responsive, and conducive to students' linguistic advancement.

In the assessment field, the incorporation of Artificial Intelligence (AI) technology has become increasingly significant in recent years. As the demand for efficient and precise language evaluation grows, traditional assessment methods often need help to keep pace with the evolving needs of students and teachers. AI-powered language

* Corresponding Author. Email: wiyaka@upgris.ac.id

© 2024 ACADEMY PUBLICATION

assessment tools offer several advantages over conventional approaches. Leveraging natural language processing, machine learning, and automated scoring algorithms, AI can provide personalized, real-time feedback to students, enabling more targeted and immediate support. By utilizing AI's ability to identify specific areas of difficulty, teachers can tailor their instructional strategies and provide personalized interventions. This allows for a more efficient and practical approach to language learning, ensuring that students receive the support they need to progress and achieve their language goals. As the field of English language education continues to evolve, the incorporation of AI-driven assessment tools and diagnostic capabilities will become increasingly crucial in providing accurate, data-driven insights and optimizing student learning outcomes.

Meanwhile, Indonesia has two critical aspects of English as a Foreign Language (EFL) education: reading anxiety and reading performance, which are due to their significant impact on students' overall language acquisition and academic success (Suryanto, 2014). Reading anxiety is a significant issue because it directly affects students' confidence and motivation when engaging with English texts. In the context of EFL in Indonesia, where English is not commonly used in daily life, students often feel intimidated by unfamiliar vocabulary, complex sentence structures, and cultural references that differ from their own (Muhlis & Fazilah, 2020). This anxiety can lead to avoidance behaviour, where students shy away from reading tasks, thus limiting their exposure to the language and hindering their progress. Additionally, reading anxiety can negatively influence students' comprehension skills, as they might struggle to focus or misunderstand the content due to stress. Ratnasari and Nurdiana (2021) found that Indonesian EFL learners reported high levels of anxiety when encountering unknown words, complex sentence structures, and culturally unfamiliar contexts in English texts. This anxiety can lead to avoidance behaviours, reduced motivation, and, ultimately, poorer reading performance.

Concurrently, reading performance among Indonesian EFL learners has concerned for educators and policymakers. The 2018 Programme for International Student Assessment (PISA) results revealed that Indonesian students' reading literacy scores were below the OECD average, ranking 72nd out of 77 participating countries/economies (OECD, 2019). In the Indonesian EFL context, strong reading performance is essential for students to excel in academic setting, where English textbooks and resources are often used. Moreover, practical reading skills enable students to access a wider range of information, thereby broadening their knowledge and improving their language competence. On the other hand, poor reading performance can limit students' academic and professional opportunities, making it a crucial area of focus in EFL education. Addressing reading anxiety and performance is essential for fostering a more effective and supportive learning environment in Indonesia. This underscores the urgent need for innovative approaches to improve English reading skills among Indonesian students (Hamra & Syatriana, 2015).

AI chatbots present a novel solution to these challenges. These digital tools can provide personalized, interactive reading experiences that may help alleviate anxiety and improve comprehension. Chatbots can offer instant vocabulary explanations, contextual examples, and comprehension checks, potentially creating a more supportive and less intimidating reading environment (Smutny & Schreiberova, 2020). Moreover, the 24/7 availability of chatbots allows for extended practice opportunities beyond the traditional classroom setting, addressing the issue of limited exposure to English in many Indonesian EFL contexts.

An increasing corpus of research has developed regarding AI chatbot-assisted language acquisition, investigating the influence of chatbots on multiple facets of language learning (Ahn, 2022; Hsu et al., 2023; Jeon, 2024; Xu et al., 2021; Ayedoun et al., 2019; Coniam, 2014). The findings of this literature indicate the advantages of AI chatbots for the advancement of foreign language acquisition (Huang et al., 2023; Lee & Hwang, 2022; Qiau & Zhao, 2023; Fryer et al., 2019). Nonetheless, further systematic investigation on AI chatbots is necessary, particularly in comparing the learning outcomes derived from interactions with chatbots against those from engaging with a teacher (Han, 2021). Moreover, the prospect of chatbots for foreign language reading remains largely unexamined; it was also found that learners utilizing AI-powered chatbots for vocabulary practice exhibited markedly superior retention rates compared to those employing conventional flashcard techniques.

Nonetheless, research particularly examining the impact of chatbots on reading anxiety and reading performance in the Indonesian EFL environment is scarce This study seeks to address this gap by examining the impact of AI chatbots on Indonesian students’ reading both anxiety and performance. This study utilized a rigorous quantitative design to evaluate an experimental group engaging in English reading through interaction with an AI-powered chatbot compared to a control group employing conventional teacher-directed approaches. This study thoroughly evaluates the efficacy of AI-based chatbot interventions compared to traditional teaching approaches.

Furthermore, this current study is driven by several compelling factors. Primarily, while chatbots have demonstrated promise in English language acquisition, their efficacy in enhancing reading skills remains insufficiently explored, with extant research yielding inconclusive outcomes. Furthermore, there is a notable paucity of studies examining AI chatbot applications among secondary school students, highlighting the imperative for additional research to corroborate and extend the generalizability of current findings. Moreover, preceding investigations into chatbot affordances have provided a theoretical foundation for exploring their potential to mitigate English reading anxiety and augment reading performance. Consequently, this study endeavours to address these identified gaps in the literature by investigating the impact of AI chatbot utilization on reading anxiety and performance among Indonesian secondary school students. To this end, the following research questions have been formulated:

© 2024 ACADEMY PUBLICATION

1. How significantly does AI chatbot-assisted assessment affect the English reading anxiety of Indonesian secondary school students?
2. How significantly does AI chatbot-assisted assessment affect the English reading performance of Indonesian secondary school students?

II. Method

A. Research Design

A quantitative design was adopted to thoroughly examine the influence of AI chatbot technology on students' English reading anxiety and performance. This approach synthesized a quasi-experimental design, allowing for the concurrent collection and analysis of quantitative to provide a holistic perspective on the phenomena under investigation.

B. Participants

The present investigation was conducted at a junior secondary institution situated in Semarang, the capital city of Central Java, Indonesia. Before the commencement of the study, ethical clearance was obtained from Universitas PGRI Semarang, as well as from the administrative body of the selected state Junior High School in Semarang. Informed consent was procured from both the student participants and their legal guardians.

Table 1
Demographic Information of Participants

Participants' information	n	Percentage (%)
Age		
13	7	10.45
14	57	85.07
15	3	4.5
Gender		
Male	24	35.82
Female	43	64.18

The study sample comprised 67 Grade 8 students, with ages ranging from 13 to 15 years, drawn from two intact classes. The participant distribution was as follows: Class A (Experiment Group) consisted of 34 students, while Class B (Control Group was composed of 33 students. These participants were enrolled in the study under the approved ethical guidelines and with full voluntary participation. While the study participants had not previously engaged with AI chatbots for English language acquisition purposes, the majority demonstrated proficiency in desktop computer utilization. These students possessed a minimum of six years of experience in English as a Foreign Language (EFL) instruction. An analysis of the participants' recent mid-term English examination results indicated that participants' English competencies were commensurate with the prediction delineated in the aforementioned curriculum.

C. Instruments

The adapted English as a Foreign Language Reading Anxiety Inventory (EFLRAI) by Zoghi (2012) was employed to assess the reading anxiety levels of the participants. The survey instrument utilized a 4-point Likert scale, with response options ranging from 1 (strongly disagree) to 4 (strongly agree). The evaluation of the improved inventory's reliability demonstrated a significant level of internal consistency, as shown by a Cronbach's alpha value of 0.921, which reflects a robust internal reliability.

To evaluate the reading abilities of the participants consistently and reliably, the study utilized the reading section of the Key English Test (KET), which is a well-regarded assessment tool for measuring the students' English proficiency. The KET reading component is designed to evaluate the proficiency of participants in comprehending fundamental written content that is frequently encountered in everyday situations. The four sections of this section, which collectively consist of 24 questions, evaluate a variety of reading abilities, such as the ability to read for gist, detail, contextual vocabulary comprehension, comprehending text structure and purpose, and drawing inferences from the provided information. This reading test was confirmed to be reliable, as evidenced by a Cronbach's alpha of 0.841, which indicates considerable internal consistency. The KET reading section was chosen for its result of validity and reliability.

D. Research Procedures

This research developed an AI chatbot called Reading Bot for student interaction via text. Accessible through a desktop webpage, Reading Bot features five pre-programmed prompts. These prompts were crafted based on Zoghi's (2012) work and the results of a pre-test English reading anxiety survey administered to participants. The prompts cover a range of topics such as grammar, cultural context, reading comprehension, and vocabularies. (see Table 2).

In conducting the study, the researchers utilized an experimental design comprising three distinct phases: pre-experimental, experimental process, and post-experimental. The pre-experimental phase involved several preparatory steps. Initially, two intact classes, presumed to possess comparable levels of reading anxiety and English reading

© 2024 ACADEMY PUBLICATION

proficiency, were identified as potential study participants. Demographic data, including self-assessed EFL competency, reading proficiency, and reading expectations, were collected from the entire sample population.

TABLE 2
EXAMPLE OF PROMPTS

Element	Example prompt
Reading performance	As an English reading assistant, your primary function is to support an Indonesian middle school student with a foundational English vocabulary of approximately 1000 words in comprehending English texts. When presented with a text in the chat interface, your responsibilities include: a. Concise Overview: Provide a brief, few-sentence summary capturing the text's central theme or main idea. b. Essential Elements: Identify and list the text's fundamental concepts or arguments using your own phrasing. c. Contemplative Inquiries: Formulate three original questions designed to encourage deeper reflection on the text and assess the student's understanding

To ensure group equivalence, particular attention was paid to establishing homogeneity between the two classes in terms of English reading anxiety and performance. The former was evaluated using a modified English reading anxiety inventory, while the latter was assessed through a Key English Test (KET) reading examination. Upon confirming the homogeneity of the two classes in both these aspects, they were randomly assigned as the experimental group (n = 34) and the control group (n = 33).

Before employing the experiment, both cohorts of study participants engaged in an introductory meeting followed by two training sessions. The experimental group members received instruction on accessing, authenticating, and engaging with the Reading Bot application via their desktop computers. This included guidance on utilizing pre-designed prompts and formulating their own inquiries. These students then spent two additional 30-minute periods practicing their interaction using reading bot, using a set of some sample texts as a basis for their exercises. Concurrently, the control group underwent a similar orientation process and participated in an equal number of reading practice sessions. These sessions utilized the identical set of four sample texts that were employed by their counterparts in the experimental group.

In conducting the experiment session, both the experiment group and control group followed a similar initial procedure. They were given 12 minutes to read two passages, each around 300 words long, and then 5 minutes to answer comprehension questions. The reading materials and questions were the same for both groups, and they were supervised by the same English teacher. The key difference was in the type of support provided after the reading task. The experiment group engaged in a 30-minute one-on-one dialogue with a chatbot to resolve any reading-related confusion. In contrast, the control group received explanations from their English teacher, who conduct reading session in a classroom setting. While the control group students could seek further clarification by interacting with the teacher and classmates, they did not have access to the Reading Bot.

The experiment was conducted over five sessions for each group. Following the experiment, all participants from both the experimental and control groups underwent the revised EFLRAI and subsequently retook a reading test from the KET. The purpose of these post-tests was to evaluate the participants' level of English reading anxiety and their reading performance after the intervention.

E. Data Analysis

This study utilized independent samples t-tests to examine the differential impact of the AI chatbot-assisted diagnostic assessment on English Reading Anxiety (between the experimental group and control group participants. A paired-samples t-test was employed to assess the pre- and post-test scores of the experimental group (EXP-G) and control group (CON-G) participants, with a particular focus on overall English reading levels. The study also employed a paired-sample t-test to investigate the effect of AI chatbot-assisted assessment usage on English reading performance. The paired-samples t-tests examined the changes in reading scores within each group.

III. FINDING

A. The Effect of AI Chatbot-Assisted Assessment on the English Reading Anxiety

To measure whether AI chatbot-assisted assessment affects the secondary students' English reading anxiety and performance, the initial investigation of their English reading anxiety was carried out. Table 2 presents the students' prior overall English readings anxiety. Point to Table 3, the Con-G exhibited a mean pre-test score of 28.5 on the English reading anxiety scale, indicating a moderate level of anxiety among its members. The scores ranged from a minimum of 20 to a maximum of 38, suggesting that while some students experienced relatively low anxiety, others faced significantly higher levels of apprehension. The standard deviation for this group was 4.25, reflecting a moderate variability in anxiety levels. This variability suggests that the Con-G comprises students with differing levels of comfort and confidence in their English reading abilities, which could influence their overall performance. In contrast, the EXP-

© 2024 ACADEMY PUBLICATION

G recorded a mean pre-test score of 32.1, which is notably higher than that of the CON-G. This indicates a higher level of anxiety regarding English reading tasks among the EXP-G participants. The scores for this group ranged from a minimum of 25 to a maximum of 40, with a standard deviation of 5.5. The increased mean score and standard deviation suggest a wider distribution of anxiety levels, indicating that some students in the EXP-G are experiencing significant anxiety that may hinder their reading performance.

TABLE 3
DESCRIPTIVE STATISTICS FOR ENGLISH READING ANXIETY IN PRE-TEST OF THE TWO GROUPS

	pre-Total English reading anxiety		pre-Average English reading anxiety		pre-General reading ability anxiety	
	CON-G	EXP-G	CON-G	EXP-G	CON-G	EXP-G
Valid	33	34	33	34	33	34
Missing	0	0	0	0	0	0
Mean	48.05	52.20	2.48	2.56	2.30	2.57
SD	13.61	12.87	0.60	0.61	0.74	0.52
Skewness	-0.52	-0.47	-0.55	-0.46	-0.36	-0.76
Std. Err of Skewness	0.38	0.38	0.38	0.38	0.38	0.38
Kurtosis	-0.61	-0.56	-0.62	-0.42	-0.93	-0.18
Std. Err of Kurtosis	0.72	0.72	0.72	0.72	0.72	0.72
Mini	21.00	22.00	1.00	1.00	1.00	1.00
Max	68.00	78.00	3.32	4.73	3.50	3.35

Meanwhile the result of post-test scores of overall English reading anxiety between the CON-G and EXP-G was presented in Table 4. Moreover, the descriptive analysis of post-test scores of overall English reading anxiety between the CON-G and the EXP-G reveal some important insights. In terms of the total English reading anxiety, the mean score for the CON-G is 48.36, with a standard deviation of 18.18, while the EXP-G has a slightly lower mean of 47.90 and a lower standard deviation of 14.62. This suggests that both groups have relatively similar levels of overall anxiety, though the EXP-G displays less variation in their anxiety levels. The skewness values for both groups are close to zero (-0.14 for CON-G and -0.034 for EXP-G), indicating that the data is fairly symmetrically distributed around the mean. Additionally, the kurtosis values of -1.08 for CON-G and -0.12 for EXP-G suggest that the distribution for the CON-G is more platykurtic, or flatter, compared to the EXP-G.

TABLE 4
DESCRIPTIVE STATISTICS FOR ENGLISH READINGS ANXIETY IN POST-TEST OF THE TWO GROUPS

	post-total English reading anxiety		post-average English reading anxiety		post- overall reading ability anxiety	
	CON-G	EXP-G	CON-G	EXP-G	CON-G	EXP-G
Valid	33	34	33	34	33	34
Missing	0	0	0	0	0	0
Mean	48.36	47.90	2.31	2.27	2.30	2.28
SD	18.18	14.62	0.88	0.72	0.87	0.75
Skewness	-0.14	-0.034	-0.150	-0.034	-0.036	-0.168
Std. Err. of	0.37	0.37	0.37	0.37-	0.37	0.37
Skewness	-1.08	-0.12	-1.07	0.12	-0.96	-0.47
Kurtosis						
Std. Err. of Kurtosis	0.74	0.74	0.74	0.74	0.74	0.74
Min	21.00	22.000	1.00	1.00	1.00	1.00
Max	84.000	86.00	4.00	4.00	4.00	4.00

Table 4 also reveals the average English reading anxiety scores. The CON-G has a mean of 2.31 with a standard deviation of 0.88, while the EXP-G has a mean of 2.27 with a lower standard deviation of 0.72. Both groups show a fairly similar average level of anxiety, though, again, the EXP-G shows less variation. The skewness and kurtosis values indicate a symmetrical distribution for both groups, with the CON-G slightly more skewed and flattened. When looking at the general reading performance anxiety, the mean for the CON-G is 2.30 with a standard deviation of 0.87, while the EXP-G shows a similar mean of 2.28 and a standard deviation of 0.75. The skewness for the CON-G is -0.036 and for the EXP-G -0.168, indicating a near-normal distribution for both groups, with the CON-G being slightly more symmetrically distributed. The kurtosis values for the CON-G and EXP-G (-0.96 and -0.47, respectively) show that the CON-G has a flatter distribution than the EXP-G.

After the pre- and post- scores of overall English readings anxiety of the two groups were measured, a paired-sample t-test was employed to see if AI chatbot-assisted assessment gives effect on English reading anxiety. These results offer

© 2024 ACADEMY PUBLICATION

a comparative analysis between the two groups, which help to determine the influence of the intervention in reducing anxiety levels related to English reading. The result of the paired samples t-test was elaborated in Table 5.

TABLE 5
DESCRIPTIVE STATISTICS FOR ENGLISH READINGS ANXIETY IN POST-TEST OF THE TWO GROUPS

Control Group (CON-G)	pre-test	post-test	*t*	*Degree for freedom*	*p*	*Effect size*	Std.err. Effect size
	pre-cumulative English reading anxiety	post-cumulative English reading anxiety	0.41	32	0.67	0.06	0.13
	pre-average English reading anxiety	post-average English reading anxiety	0.49	32	0.68	0.07	0.15
	pre-overall reading ability anxiety	post-overall reading ability anxiety	0.21	32	0.89	0.04	0.15
Experiment Group (EXP-G)	pre-cumulative English reading anxiety	post-cumulative English reading anxiety	2.36	33	0.03	0.42	0.16
	pre-average English reading anxiety	post-average English reading anxiety	2.52	33	0.03	0.43	0.17
	pre-overall reading ability anxiety	post-overall reading ability anxiety	2.31	33	0.03	0.36	0.15

Based on Table 5, the results for the CON-G indicate no significant differences between the pre-test and post-test scores for any of the three measures of English reading anxiety. For total English reading anxiety, the t-value is 0.41 with a p-value of 0.67, indicating that the difference between pre-test (M = 48.36) and post-test scores is statistically insignificant. The effect size, represented by Cohen's d (0.06), is very small, suggesting minimal practical significance in the observed change. Similarly, for average English reading anxiety, the t-value is 0.49 with a p-value of 0.68, showing that there is no meaningful reduction in anxiety after the control period. The effect size is again very small (Cohen's d = 0.07), reinforcing the conclusion that no substantial change occurred. In terms of general reading ability anxiety, the CON-G also shows no significant difference between pre-test and post-test scores (t = 0.21, p = 0.89). With Cohen's d calculated at 0.04, the effect size is negligible, indicating virtually no impact on reading ability anxiety in the CON-G.

In contrast, the results for the EXP-G demonstrate statistically significant improvements in reducing English reading anxiety. For total English reading anxiety, the t-value is 2.36 with a p-value of 0.03, suggesting a significant reduction in anxiety levels following the intervention (AI Chatbot-Assisted Assessment). The effect size (Cohen's d = 0.42) indicates a moderate practical impact, highlighting that AI chatbot-assisted assessment had a meaningful effect on reducing overall English reading anxiety. For average English reading anxiety, the paired samples t-test also reveals a significant decrease from pre-test to post-test scores (t = 2.52, p = 0.03). Cohen's d is 0.43, which again suggests a moderate effect size, indicating that the intervention was effective in reducing average levels of English reading anxiety. Finally, the analysis of general reading ability anxiety in the EXP-G shows a significant reduction, with a t-value of 2.31 and a p-value of 0.03. The effect size (Cohen's d = 0.36) is slightly smaller compared to the other anxiety measures but still suggests a moderate impact of AI chatbot-assisted assessment.

The results of the paired samples t-tests highlight a clear distinction between the CON-G and the EXP-G. While the CON-G showed no significant changes in any of the anxiety measures, the EXP-G exhibited statistically significant reductions in total, average, and general reading ability anxiety. These findings suggest that the intervention applied to the EXP-G was effective in lowering English reading anxiety, with moderate effect sizes across all measures. The lack of significant change in the CON-G underscores the importance of AI chatbot-assisted assessment in mitigating anxiety levels related to English reading.

B. The Effect of AI Chatbot-Assisted Assessment on the English Reading Performance

This second section explores the role of artificial intelligence in enhancing students' reading abilities through interactive assessments. As educational technologies evolve, AI chatbots have gained attention for their potential to offer personalized feedback, engage learners in real-time, and support teachers in monitoring progress. This section examines the impact of implementing AI chatbot-assisted assessments on English reading performance, focusing on how these tools influence students' reading comprehension and foster a more engaging learning environment in secondary schools.

© 2024 ACADEMY PUBLICATION

TABLE 6
DESCRIPTIVE STATISTICS FOR ENGLISH READING PERFORMANCE OF THE TWO GROUPS

	pre-Reading	post-Reading	pre-Reading	post-Reading
	CON-G	CON-G	EXP-G	EXP-G
Valid	33	33	34	34
Missing	0	0	0	0
Mean	15.72	16.78	15.92	17.86
SD	4.61	4.95	4.68	5.20
Skewness	-0.78	-0.78	-0.46	-0.35
Std. Err. of Skewness	0.37	0.37	0.37	0.37
Kurtosis	-0.57	0.24	-0.99	-1.02
Std. Err. of Kurtosis	0.76	0.76	0.76	0.76
Min	5.00	4.00	6.00	6.00
Max	23.000	25.000	24.000	26.000

The analysis of the effect of AI chatbot-assisted assessment on the students' English reading performance was first elaborated in Table 6 showing the result of pre- and post-test of students' reading performance through KET. The descriptive statistics in Table 4 reveal notable differences between the CON-G and the EXP-G. Before the intervention, the mean reading score for the CON-G was 15.72 (SD = 4.61), while the EXP-G had a slightly higher mean score of 15.92 (SD = 4.68). After the intervention, both groups demonstrated improvement in their reading performance, but the EXP-G showed a more significant increase. The CON-G's mean post-test score rose to 16.78 (SD = 4.95), while the EXP-G's post-test score reached 17.86 (SD = 5.20), indicating a greater improvement in the EXP-G compared to the CON-G. The skewness and kurtosis values for both groups suggest a generally normal distribution of reading scores, with slight variations in the shape of the distributions. Overall, these results suggest that AI chatbot-assisted assessments had a more substantial impact on enhancing reading performance in the EXP-G than in the CON-G.

TABLE 7
PAIRED SAMPLES T-TESTS OF ENGLISH READING PERFORMANCE WITHIN THE TWO GROUPS

	Pre-test	Post-test	*t*	*Degree for freedom*	*p*	*Effect size*	Std.err. Effect size
CON-G	Pre-reading	Post-reading	-1.03	32	0.33	-0.16	0.13
EXP-G	Pre-reading	Post-reading	-2.83	33	0.04	-0.82	0.12

The paired samples t-tests in Table 7 evaluates changes in English reading performance between pre- and post-test scores for both the control group (CON-G) and experimental group (EXP-G). Results indicate a significant improvement in the experimental group, while the control group showed minimal change. The control group demonstrated a slight enhancement in reading performance ($t(32) = -1.03$, $p = 0.33$, $d = -0.16$), but this change was not statistically significant. The small effect size (Cohen's $d = -0.16$) suggests minimal practical impact, with a standard error of 0.13 indicating reasonable precision in this estimate. In contrast, the experimental group exhibited a substantial and statistically significant improvement ($t(33) = -2.83$, $p = 0.04$, $d = -0.82$). The p-value of 0.04 falls below the conventional threshold of 0.05, indicating statistical significance. The large effect size (Cohen's $d = -0.82$) suggests a strong practical significance of the intervention, with the low standard error (SE = 0.12) providing high confidence in this estimate. The stark difference between the CON-G and EXP-G results strongly suggests that the intervention had a powerful and meaningful impact on improving English reading performance in the experimental group. These findings provide robust evidence for the efficacy of the intervention, demonstrating both statistical significance and substantial practical importance in enhancing students' English reading performance.

IV. DISCUSSION

The implementation of AI chatbot-supported assessment in educational environments has attracted interest due to its capacity to decrease anxiety and improve reading proficiency in students. This discussion is on the impact of AI chatbot-supported assessment on English reading anxiety and reading performance among secondary school students in Indonesia. It compares an experimental group that deployed AI chatbot-assisted assessment with a control group that did not receive this intervention.

AI chatbots facilitated by sophisticated natural language processing technologies can encourage students with them. For instance, this AI technology for English reading practice could enable students to participate in the chatbot's interactive reading activity and receive explanations and reading strategies for unknown words. Thus, the anxiety scores dropped significantly among the experimental group monitored with AI chatbot-assisted assessments. The reading endeavour in the group also became a less stressful experience because, unlike in physical classroom snapped due to teachers, learners could easily navigate in task discussion, and allay teacher wrath that rang the bell of anxiety (El Shazly, 2021).

In contrast, the control group, who did not get AI chatbot-assisted assessment, exhibited no any notable alterations in their reading anxiety or performance. The absence of assistance resulted in students being subjected to the customary

© 2024 ACADEMY PUBLICATION

pressures linked to reading assessments, such as the anxiety caused by timed tests and the possibility of receiving negative comments from both peers and instructors (Sumakul et al., 2022; Alshumaimeri & Alshememry, 2023). The lack of individualized assistance undoubtedly played a role in an ongoing rise of anxiety levels among these students (Moybeka et al., 2023).

In addition, the experimental group not only had a decrease in anxiety but also showed enhanced proficiency in reading. The AI chatbot offered customized reading materials that corresponded to the students' aptitude levels, allowing them to interact with texts that were yet demanding and approachable (Jiang, 2022). The implementation of this adaptive learning technique enabled students to develop self-assurance in their reading skills, resulting in enhanced performance on examinations. The chatbot's role in providing a low-pressure setting for practicing reading skills was essential for improving their overall reading proficiency (Tang & Foley, 2022). Conversely, the control group's reading performance did not improve. The absence of instant feedback and adaptive learning possibilities hindered students' engagement with reading materials. The conventional evaluation techniques utilized in the control group failed to address the specific learning requirements of each individual, a crucial factor in promoting advancements in language proficiency (Alhalangy & AbdAlgane, 2023). The lack of progress in this area emphasizes the significance of incorporating technology into education in order to address learners' varying demands.

Furthermore, the AI chatbot's ability to offer round-the-clock assistance was a notable benefit for the experimental group. Students can engage with the chatbot beyond regular school hours, enabling them to further practice and strengthen their abilities at their own preferred speed. This adaptability is especially advantageous for students who may struggle to manage their time or have additional obligations beyond their academic pursuits. The control group, however, did not have the chance to engage in extended learning, which could have influenced their performance levels to remain the same.

The results of this study are consistent with prior research that underscores the significance of technology in alleviating language learning anxiety. Research has demonstrated that utilizing artificial intelligence (AI) in language teaching can establish a more nurturing learning atmosphere, which is crucial for students who encounter anxiety associated with language acquisition (Zheng, 2024; Rukiati et al., 2023). The favourable effects observed in the experimental group provide further evidence that technology can effectively be used as a beneficial tool to alleviate emotional obstacles to learning.

Furthermore, the societal dimension of acquiring knowledge via AI chatbots should not be disregarded. The students in the experimental group expressed a stronger sense of connection to their learning experience, since the chatbot enabled a conversation that simulated exchanges with a helpful peer or instructor. The feeling of connection is crucial for fostering motivation and engagement, both of which are key elements in achieving success in language acquisition (Lee & Hwang, 2022). Conversely, the control group did not have this participatory component, which could have perhaps led to their sense of seclusion and unease.

V. Conclusion

In conclusion, the implementation of AI chatbot-assisted assessments has profound implications for reducing English reading anxiety and improving reading performance among Indonesian secondary school students. The experimental group benefited significantly from the personalized, adaptive learning environment created by the chatbot, while the control group remained unchanged due to the absence of such interventions. AI chatbots integration provides a customized and dynamic learning experience that specifically tackles the distinct difficulties encountered by EFL learners, such as new terminology and cultural situations. Chatbots promote increased participation and self-assurance in language usage by offering prompt feedback and creating an atmosphere. This research emphasizes the necessity of additional investigation into the use of AI technology in language learning inside educational institutions. It advocates the integration of novel tools to enhance student achievement in English literacy.

These findings suggest that educational institutions should consider integrating AI technologies into their curricula to enhance student learning experiences and outcomes. The results of this study contribute to the growing body of literature on the effectiveness of AI in education, particularly in language learning contexts. Future research should explore the long-term effects of AI chatbot-assisted learning on various aspects of language acquisition and consider the implications for broader educational practices.

Acknowledgements

This research has been funded by Indonesian Ministry of Education, Culture, Research and Technology through Regular Fundamental Research Grant No108/E5/PG.02.00.PL/2024.

References

[1] Ahn, S. (2022). The effects of chatbot on grammar competence for Korean EFL college students. *Journal of Digital Convergence*, *20*(3), 53-61.

[2] Alhalangy, A., & Abdalgane, M. (2023). Exploring the impact of ai on the EFL context: a case study of Saudi universities. *Journal of Intercultural Communication*, *23*(2), 41-49. https://doi.org/10.36923/jicc.v23i2.125

© 2024 ACADEMY PUBLICATION

[3] Alshumaimeri, Y. A., & Alshememry, A. K. (2023). The extent of ai applications in EFL learning and teaching. *IEEE Transactions on Learning Technologies*. https://doi.org/10.1109/TLT.2023.3322128
[4] Ayedoun, E., Hayashi, Y., & Seta, K. (2019). A conversational agent to encourage willingness to communicate in the context of English as a foreign language. *Journal of Educational Computing Research*, *57*(3), 555-580.
[5] Bibauw, S., François, T., & Desmet, P. (2019). Discussing with a computer to practice a foreign language: Research synthesis and conceptual framework of dialogue-based CALL. *Computer Assisted Language Learning*, *32*(8), 827-877. https://doi.org/10.1080/09588221.2018.1535508
[6] Chapelle, C. A., & Sauro, S. (Eds.). (2017). *The Handbook of Technology and Second Language Teaching and Learning.* John Wiley & Sons. https://doi.org/10.1002/9781118914069
[7] Coniam, D. (2014). The linguistic accuracy of chatbots: usability from an ESL perspective. *Text & Talk*, *34*(5), 545-567.
[8] El Shazly, R. (2021). Effects of artificial intelligence on English speaking anxiety and speaking performance: A case study. *Expert Systems*, *38*(3), e12667. https://doi.org/10.1515/text-2014-0018
[9] Fryer, L. K., Nakao, K., & Thompson, A. (2019). Chatbot learning partners: connecting learning experiences, interest and competence. *Computers in Human Behavior*, *93*, 279-289.
[10] Jeon, J. (2024). Exploring ai chatbot affordances in the EFL classroom: young learners' experiences and perspectives. *Computer Assisted Language Learning*, *37*(1-2), 1-26.
[11] Hamra, A., & Syatriana, E. (2015). Developing a model of teaching reading comprehension for EFL students. *TEFLIN Journal*, *21*(1), 27-4.
[12] Han, D. (2021). An analysis of Korean EFL learners' experience on English classes using ai chatbot. *Robotics & AI Ethics*, *6*(3), 1-9.0.
[13] Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.
[14] Hsu, T. C., Chang, C., & Jen, T. H. (2023). Artificial intelligence image recognition using self-regulation learning strategies: effects on vocabulary acquisition, learning anxiety, and learning behaviors of English language learners. *Interactive Learning Environments*, 1-19. https://doi.org/10.1080/10494820.2023.2165508
[15] Huang, X., Zou, D., Cheng, G., Chen, X., & Xie, H. (2023). Trends, research issues and applications of artificial intelligence in language education. *Educational Technology & Society*, *26*(1), 112-131.
[16] Jiang, R. (2022). How does artificial intelligence empower EFL teaching and learning nowadays? a review on artificial intelligence in the EFL context. *Frontiers in Psychology*, *13*, 1049401. https://doi.org/10.3389/fpsyg.2022.1049401
[17] Lauder, A. (2008). The status and function of English in Indonesia: a review of key factors. *Makara Human Behavior Studies in Asia*, *12*(1), 9-20. https://doi.org/10.7454/mssh.v12i1.128
[18] Lee, J. Y., & Hwang, Y. (2022). A meta-analysis of the effects of using ai chatbot in Korean EFL education. *Studies in English Language & Literature*, *48*(1), 213-243. https://doi.org/10.21087/nsell.2022.11.83.213
[19] Moybeka, A. M., Syariatin, N., Tatipang, D. P., Mushthoza, D. A., Dewi, N. P. J. L., & Tineh, S. (2023). Artificial intelligence and English classroom: the implications of ai toward EFL students' motivation. *Edumaspul: Jurnal Pendidikan*, *7*(2), 2444-2454.
[20] Muhlis, A., & Fazilah, N. (2020). Foreign language reading anxiety among Indonesian EFL learners. *Journal of Research in Foreign Language Teaching*, *2*(1), 23-36. https://doi.org/10.33487/edumaspul.v7i2.6669
[21] OECD. (2019). PISA 2018 Results (Volume I): *What Students Know and Can Do.* OECD Publishing, Paris.
[22] Ratnasari, D., & Nurdiana, N. (2021). Investigating EFL students' reading anxiety: causes and strategies. *Journal of English Language Teaching and Linguistics*, *6*(1), 155-170.
[23] Rukiati, E., Wicaksono, J. A., Taufan, G. T., & Suharsono, D. D. (2023). AI on learning English: application, benefit, and threat. *Journal of Language, Communication, and Tourism*, *1*(2), 32-40. https://doi.org/10.25047/jlct.v1i2.3967
[24] Smutny, P., & Schreiberova, P. (2020). Chatbots for learning: a review of educational chatbots for the Facebook Messenger. *Computers & Education*, *151*, 103862. https://doi.org/10.1016/j.compedu.2020.103862
[25] Sumakul, D. T. Y., Hamied, F. A., & Sukyadi, D. (2022). Artificial intelligence in EFL classrooms: friend or foe?. *LEARN Journal: Language Education and Acquisition Research Network*, *15*(1), 232-256.
[26] Suryanto, S. (2014). Issues in teaching English in a cultural context: a case of Indonesia. *Journal of English Literacy Education*, *1*(2), 75-82.
[27] Tang, J., & Foley, J. (2022). A case study on the effectiveness of applying content and language integrated learning in an artificial intelligence English reading course. *Arab World English Journal*, *13*(3). https://doi.org/10.24093/awej/vol13no3.15
[28] Qiao, H., & Zhao, A. (2023). Artificial intelligence-based language learning: illuminating the impact on speaking skills and self-regulation in Chinese EFL context. *Frontiers in Psychology*, *14*, 1255594. https://doi.org/10.3389/fpsyg.2023.1255594
[29] Xu, Z., Chen, Z., Eutsler, L., Geng, Z., & Kogut, A. (2021). A scoping review of digital game-based technology on English language learning. *Educational Technology Research and Development*, *68*(3), 877-904. https://doi.org/10.1007/s11423-019-09702-2
[30] Zheng, S. (2024). The effects of chatbot use on foreign language reading anxiety and reading performance among Chinese secondary school students. *Computers and Education: Artificial Intelligence*, *7*, 100271. https://doi.org/10.1016/j.caeai.2024.100271

Wiyaka is an associate professor in English Department of Universitas PGRI Semarang. He earned his doctoral degree (Dr.) in Language Education from Universitas Negeri Semarang, Indonesia. In addition to his teaching responsibilities, Dr. Wiyaka has been actively involved in multiple research initiatives funded by the Ministry of Higher Education and various external sponsors. His academic pursuits and research interests are primarily focused on language assessment methodologies, reading pedagogy, and writing

© 2024 ACADEMY PUBLICATION

instruction. Dr. Wiyaka has made significant contributions to the field through the publication of research articles in peer-reviewed journals and has authored authoritative textbooks on assessment practices.

Lusia Maryani Silitonga holds a doctoral degree from the Department of Technological and Vocational Education at the National Yunlin University of Science and Technology. Her research focuses on applying EEG in education, game-based learning in entrepreneurship education, AI-based education, and computational thinking in academic writing. Dr. Silitonga's work combines neuroscience, educational technology, and computer science to develop innovative higher education approaches.

Sunardi is a senior lecturer in applied linguistics at Universitas Dian Nuswantoro's English Department in Semarang, Indonesia. He earned his doctorate in linguistics from Universitas Sebelas Maret Surakarta, focusing on English lecture discourse in Indonesian higher education using systemic functional linguistics. His research interests include classroom discourse analysis and inclusive learning. Dr. Sunardi has presented and published his work in national and international conferences and scholarly journals.

Yuventius Tyas Catur Pramudi teaches some courses in information system at the department of information system in Universitas Dian Nuswantoro, Semarang Indonesia. He completed his doctoral degree in economics at Universitas Merdeka Malang, Indonesia. Some research has been conducted by him that cover many interdisciplinary studies: information technology supports for business, education, and arts. He has participated and published in many national and international seminars and journals.

© 2024 ACADEMY PUBLICATION