Linguistic, English Education and Art (LEEA) Journal
Volume 9 Nomor 2, Januari-April 2026
e-ISSN : 2597-3819
p-ISSN : 2597-9248
DOI : https://doi.org/10.31539/9vr4j332

LEVERAGING EDBOT.AI AS A DIAGNOSTIC ASSESSMENT TOOL TO ENHANCE THE STUDENTS' ENGLISH ACHIEVEMENT

Alya Shafa Nabila[1]
Universitas PGRI Semarang

Entika Fani Prastikawati[2]
Universitas PGRI Semarang

Maria Yosephin Widarti Lestari[3]
Universitas PGRI Semarang

alyashfnbl13@gmail.com[1]

Submit, 05-03-2026 *Accepted*, 26-04-2026 *Publish*, 27-04-2026

ABSTRACT

The study aims to examine the effect of Edbot.ai on students' English achievement and to investigate students' perceptions of its implementation. To achieve these aims, the study employed a sequential explanatory mixed-methods design using a quasi-experimental approach. To get the data, the study used English test and questionnaire. The findings reveal that the use of Edbot.ai contributed positively to students' English achievement. Students also perceived Edbot.ai as supportive, engaging, and helpful in understanding their learning progress through timely feedback. Overall, the study indicates that AI-based diagnostic assessment can effectively support English learning outcomes. Future research is recommended to investigate the long-term impact of Edbot.ai to deepen understanding of its instructional potential.

Keywords: Edbot.ai, Diagnostic Assessment, English Achievement, Students' Perception

INTRODUCTION

Assessment plays a critical role in educational progress, serving not only as a benchmark for student learning but as an instrument that diagnoses gaps and guides instructional remediation. In Indonesia, the conventional assessment model has largely focused on summative measures final exams and standardized tests that fall short in uncovering the nuanced learning difficulties encountered by students (Tosuncuoglu, 2018; Wiyaka et al., 2025). Yet, under the current curriculum, educators are encouraged to embrace more diagnostic approaches that

inform personalized learning toward mastery (Maulidyah, 2025). Diagnostic assessment helps teachers pinpoint learners' strengths and weaknesses be it in grammar, reading comprehension, or oral skills enabling more targeted support and thereby fostering deeper understanding and learner confidence (Jimola & Ofodu, 2019). Without such diagnostic clarity, instructional strategies tend to remain uniform and may inadvertently neglect persistent language learning gaps (Farhadi & Selcuk, 2025). This shortfall underscores the urgent need for assessment tools that can deliver specific, actionable insights into student learning trajectories.

Across Indonesia, a handful of studies affirm the value of diagnostic assessments in EFL settings. For example, using diagnostic techniques in teaching explanation texts led to a measurable improvement in English competence (Widiyanti, et al., 2022). In another case, junior high classes benefited when a diagnostic test was administered before instruction, enabling teachers to better tailor their lesson content (Aditama et al., 2023). Although mostly manual or low-tech, these interventions highlight the potential power of diagnostic tools in guiding meaningful English instruction. Other research in local language instruction and multimodal settings further demonstrates how individualized feedback, predicated on diagnostic analysis, supports receptive skill development (Amin, 2023). Collectively, these studies illuminate the pedagogical advantages of diagnostic approaches, while simultaneously revealing that their routine adoptio, especially in digital form remains limited in Indonesian classrooms.

At the same time, Indonesia is experiencing a gradual shift toward integrating educational technology into language instruction. Tools like Google Forms, Kahoot!, and Jamboard are increasingly embraced for formative assessments and student engagement (Nuraini et al., 2025; Setiyowati & Ardaniah, 2025). While teachers show favorable attitudes toward ed-tech (e.g., average acceptance ratings of around 4.2 out of 5), actual implementation is moderate, often hindered by infrastructure, training, or accessibility issues (Koraishi, 2023; Lee at al., 2023). Despite challenges, technology has unlocked new possibilities for individualized instruction, delivering instant feedback, enabling real-time monitoring, and accommodating diverse learning styles. Remote and blended learning environments, particularly during and after the pandemic, have reinforced how digital platforms can sustain engagement and practice in areas like reading, pronunciation, and vocabulary acquisition (Fang et al., 2023). However, existing tools generally lack AI-driven diagnostics customized for junior high EFL learners, signaling both a need and an opportunity for more advanced, adaptive technologies.

Edbot.ai as a pioneering AI-powered learning coachn is crafted by Solve Education! Too democratize educational access and personalization. Edbot.ai

functions as an interactive chatbot accessible via a web interface, designed around gamified experiences, adaptive content, and a robust rewards system (Luo et al., 2025; Mohammadi, 2025). It delivers structured learning paths in subjects such as English, math, and life skills, complemented by real-time feedback and performance assessments. Learners can earn points, redeem rewards, and build community connections, all of which motivate sustained engagement. The platform's framework focusing on Gamification, AI Coach, Incentives, and Network creates an ecosystem of personalized learning that scales across diverse learner populations.

One of Edbot.ai's most compelling features is its AI-powered "Chat" capability for English practice, providing learners with instant in-chat corrections of grammar, vocabulary, spelling, and phrasing. This immersive experience facilitates low-stakes, high-frequency learning and fosters meaningful communicative practice (González-Calatayud et al., 2021). Moreover, its commitment to accessibility is seen in partnerships like ecopower, which brings inclusive learning to persons with disabilities in Indonesia is evident in Edbot.ai's design (Swiecki et al., 2022). The platform supports low-bandwidth access via channels like Telegram, ensuring learners in remote or underserved communities can engage with quality learning resources. Together, these elements position Edbot.ai not only as a tool for learning, but also as a scalable, inclusive solution for adaptive instruction in diverse environments.

In Indonesian EFL contexts, AI-powered diagnostic tools like Edbot.ai remain underexplored. Although AI chatbots have been shown to improve vocabulary mastery in junior high settings elsewhere in Indonesia, such studies typically use rudimentary platforms rather than integrated, gamified, adaptive systems (Lui & Bui, 2025; Yang, 2025; Taj et al., 2025). The combination of automated diagnostic assessment, instant corrective feedback, gamified motivation, and accessibility represented by Edbot.ai has yet to be systematically investigated in junior high English classrooms. This gap highlights the need to evaluate both how such a tool affects learning outcomes and how students respond to its use, particularly considering the developmental and motivational needs of adolescents transitioning to higher-level English language proficiency.

This study addresses this gap by exploring Edbot.ai's role as a diagnostic assessment tool in junior high English education in Indonesia. By examining both objective learning outcomes and subjective perceptions, the research seeks to determine whether Edbot.ai can enhance English achievement and how learners experience its integration be it in terms of engagement, ease of use, motivational impact, or perceived relevance. The study's dual focus aligns with pedagogical imperatives for both efficacy and learner voice in educational technology adoption. Accordingly, the study is guided by two central questions: (1) How does

Edbot.ai as diagnostic assessment affect the students' English achievement? and (2) How is Edbot.ai as diagnostic assessment perceived by the students in junior high school?

LITERATURE REVIEW

Assessment has long been recognized as a central component in the educational process, functioning not only to evaluate learning outcomes but also to inform instructional practices. Within the field of English as a Foreign Language (EFL), assessment is increasingly viewed as a dynamic tool that can diagnose learners' needs and guide pedagogical decisions. Traditional assessment practices in Indonesia, however, have predominantly emphasized summative approaches such as final examinations and standardized testing. These methods often fail to capture the complexity of students' learning difficulties, particularly in language acquisition, where skills such as grammar, reading comprehension, and speaking require continuous monitoring and feedback (Tosuncuoglu, 2018; Wiyaka et al., 2025). In response to these limitations, recent curricular reforms advocate for more diagnostic and formative forms of assessment that support personalized and mastery-oriented learning (Maulidyah, 2025).

Diagnostic assessment, in particular, has gained attention for its ability to provide detailed insights into students' strengths and weaknesses. According to Jimola and Ofodu (2019), diagnostic assessment enables teachers to identify specific areas of difficulty and tailor instruction accordingly, thereby enhancing both learning outcomes and student confidence. Without such targeted insights, teaching strategies tend to remain generalized and may overlook persistent learning gaps (Farhadi & Selcuk, 2025). This underscores the importance of adopting assessment tools that offer actionable and individualized feedback, especially in EFL contexts where learners exhibit diverse proficiency levels.

Empirical studies conducted in Indonesia further support the effectiveness of diagnostic assessment in improving English language learning. Widiyanti et al. (2022) found that the use of diagnostic techniques in teaching explanation texts led to significant improvements in students' English competence. Similarly, Aditama et al. (2023) demonstrated that administering diagnostic tests prior to instruction allowed teachers to design more targeted lesson plans, resulting in better student engagement and comprehension. In addition, Amin (2023) highlighted the role of diagnostic-based individualized feedback in enhancing receptive skills within multimodal learning environments. Despite these promising findings, most implementations remain manual or low-tech, indicating that the integration of more advanced and scalable diagnostic tools is still limited in Indonesian classrooms.

Parallel to the growing emphasis on diagnostic assessment, there has been a gradual integration of educational technology in Indonesian education. Digital tools such as Google Forms, Kahoot!, and Jamboard are increasingly utilized to facilitate formative assessment and interactive learning experiences (Nuraini et al., 2025; Setiyowati & Ardaniah, 2025). Research indicates that teachers generally hold positive attitudes toward educational technology, with relatively high acceptance levels (Koraishi, 2023; Lee et al., 2023). However, actual implementation remains moderate due to challenges such as limited infrastructure, insufficient training, and unequal access. Nonetheless, technology has expanded opportunities for personalized learning by enabling real-time feedback, adaptive content delivery, and flexible learning environments. This became particularly evident during the shift to remote and blended learning, where digital platforms played a crucial role in maintaining student engagement and supporting language skill development (Fang et al., 2023).

Despite these advancements, most existing digital tools lack sophisticated AI-driven diagnostic capabilities tailored to EFL learners, especially at the junior high school level. This limitation points to the need for more innovative solutions that combine diagnostic assessment with adaptive learning technologies. One such emerging platform is Edbot.ai, an AI-powered learning coach developed by Solve Education! Edbot.ai integrates gamification, artificial intelligence, and personalized learning pathways to create an engaging and adaptive learning environment. The platform offers structured content in English and other subjects, along with real-time feedback and performance tracking. Its reward system and community-based features further enhance learner motivation and sustained engagement (Luo et al., 2025; Mohammadi, 2025).

A notable feature of Edbot.ai is its AI-driven chat function, which allows learners to practice English interactively while receiving immediate corrections in grammar, vocabulary, and sentence structure. This form of low-pressure, high-frequency interaction supports communicative competence and encourages active language use (González-Calatayud et al., 2021). Additionally, the platform emphasizes accessibility by supporting low-bandwidth environments and integrating with widely used communication channels such as Telegram. Partnerships aimed at inclusive education, including support for learners with disabilities, further demonstrate its potential as a scalable and equitable learning solution (Swiecki et al., 2022).

Nevertheless, the application of AI-powered diagnostic tools like Edbot.ai in Indonesian EFL classrooms remains underexplored. While previous studies have shown that chatbot-based learning can improve vocabulary mastery, these studies often rely on simpler, less integrated technologies (Lui & Bui, 2025; Yang, 2025; Taj et al., 2025). The unique combination of automated diagnostic

assessment, real-time feedback, gamification, and accessibility offered by Edbot.ai has not yet been systematically investigated, particularly among junior high school learners. This gap highlights the need for further research to evaluate both the effectiveness of such tools in improving learning outcomes and students' perceptions of their use. Understanding these aspects is crucial, especially given the developmental characteristics and motivational needs of adolescent learners transitioning to higher levels of English proficiency.

RESEARCH METHOD

This study employed a sequential explanatory mixed-method design, in which the quantitative phase was prioritized, followed by the qualitative phase to provide deeper insights into the results (Creswell, 2018). In line with this, the present research applied the design not only to measure the effectiveness of the intervention statistically but also to explore students' subjective experiences that could not be fully captured through numerical data. The quantitative phase adopted a quasi-experimental design involving two intact classes that were purposively assigned as the experimental and control classes. After the quantitative data were analyzed and the statistical results obtained, the qualitative phase was conducted to explain, elaborate, and enrich the findings. This phase focused on capturing students' perceptions, challenges, and attitudes toward the use of Edbot.ai, thereby providing a more comprehensive understanding of its role in diagnostic assessment.

This study involved 64 seventh-grade students from SMP Negeri 12 Semarang. The participants were chosen using purposive sampling to ensure that the selected classes had comparable levels of English proficiency, thereby allowing for balanced group comparisons. Table 1 summarizes the characteristics of the participants in terms of group and gender distribution. In the control class, there were 16 male and 14 female students, whereas the experimental class comprised 8 male and 22 female students. Furthermore, seventh grade was intentionally selected because diagnostic assessment aims to identify areas in which learners need further development (Brown, 2019). Conducting the study at the beginning of junior high school allows the diagnostic outcomes to provide a basis for enhancing teaching and learning in the subsequent grades.

Table 1. Information on Participants in The Study

Group	Items	N
Control Class	Male	16
	Female	14
Experimental Class	Male	8
	Female	22

To collect both quantitative and qualitative data, four research instruments were employed in this study, namely a pre-test, a diagnostic assessment, a post-test, and an open-ended questionnaire. Each instrument was designed to align with the research objectives and to ensure the validity of the findings.

The pre-test was administered before the intervention to measure students' initial achievement in English. It was designed in the form of an integrated classroom test, similar to regular school examinations, and consisted of a combination of tasks rather than relying solely on multiple-choice questions. The results served as a baseline to compare with students' later performance.

The diagnostic assessment phase consisted of two different instruments: (a) Edbot.ai-based assessment and (b) a traditional paper-based diagnostic test. Both instruments were developed based on the same indicators and assessment criteria, covering speaking, listening, reading, and grammar. The experimental class was assessed using Edbot.ai, while the control class used the manual test. This design enabled a direct comparison of the effectiveness of the technology-based tool and the traditional approach. After the intervention, a post-test was administered with a format similar to the pre-test. The test functioned to evaluate students' achievement after being exposed to either Edbot.ai or the traditional diagnostic assessment. The comparison of pre-test and post-test results provided quantitative evidence of students' progress.

To obtain qualitative data, an open-ended questionnaire was distributed to students in the experimental class. The questionnaire focused on their experiences, perceptions, and challenges in using Edbot.ai as a diagnostic assessment tool. The responses were analyzed to complement the quantitative results, offering a deeper explanation and contextual understanding of the findings.

The data collection procedures in this study were carried out in several sequential stages, involving both quantitative and qualitative strands to align with the explanatory mixed-method design. At the beginning of the study, a pre-test was administered to both the experimental and control classes during regular class hours under the supervision of the teacher to ensure uniform conditions across the groups. During the intervention phase, the diagnostic assessment was implemented, with the experimental class completing their diagnostic tasks through the Edbot.ai platform and the control class using a traditional paper-based format. Both groups completed the assessment within the same time frame, monitored by the researcher and teacher to maintain fairness and consistency. After the intervention, a post-test was conducted for

both groups in the same setting and under the same procedures as the pre-test, ensuring comparability between the two data sets. In addition, qualitative data were collected from the experimental class using an open-ended questionnaire distributed via Google Form, which students completed outside of class hours to allow sufficient time for reflection on their experiences and perceptions of using Edbot.ai as a diagnostic assessment tool. Through this systematic sequence, the study gathered both quantitative evidence of students' achievement and qualitative insights into their experiences, ensuring that the data collection process comprehensively addressed the research objectives.

The quantitative data obtained from the pre-test and post-test were analyzed using SPSS to measure the differences between the experimental and control classes. Descriptive statistics were used to present the mean scores and standard deviations, while inferential statistics, specifically the Independent Samples t-test and N-gain score analysis, were employed to determine the significance of differences in learning achievement. The Independent Samples t-test was selected because it is appropriate for comparing the mean post-test scores of two independent groups, namely the experimental class and the control class, to examine whether the use of Edbot.ai resulted in significantly different learning outcomes compared to conventional written diagnostic assessment. Meanwhile, the N-gain score analysis was used to measure the magnitude of learning improvement from pre-test to post-test and to compare the effectiveness of the intervention between the two groups. The qualitative data, derived from open-ended questionnaires, were analyzed using thematic analysis. This technique was chosen because it allows for the identification of recurring themes, patterns, and categories within students' responses, making it suitable for capturing the nuanced perceptions and experiences of learners (Braun & Clarke, 2006). The integration of both data strands enabled the researcher to draw comprehensive conclusions about both the effectiveness of Edbot.ai in improving English achievement and the students' perceptions of its use as a diagnostic tool.

FINDING

The Effect of Edbot.ai as a Diagnostic Assessment Tool on Students' English Achievement

To analyze the effect of Edbot.ai on students' English achievement, a quantitative analysis was conducted using an Independent Samples t-test and N-Gain score analysis. The analysis focused on the pre-test and post-test scores of both the control and experimental class. The control class consisted of students who were assessed using traditional written diagnostic test, while the experimental class consisted of students who were assessed using Edbot.ai as a diagnostic assessment tool. A descriptive statistical analysis was first carried out

to provide an overview of students' pre-test and post-test scores in both the control and experimental class. The results of descriptive statistical analysis are presented in Table 2.

Table 2. Descriptive statistical analysis of control and experimental class's pre-test and post-test

Statistics	Control Class		Experimental Class	
	Pre-Test	**Post-Test**	**Pre-Test**	**Post-Test**
N (Sample total)	30	30	30	30
Mean	64.00	65.00	71.16	86.5
Median	60.00	65.00	70.00	85.00
Std. Deviation	16.84	18.71	7.39	7.56
Minimum Score	25	25	55	70
Maximum Score	100	100	90	100

Descriptive statistical analysis shows that both groups experienced improvement in their post-test scores. However, the improvement in the experimental class was substantially greater than that of the control class. The mean score of the experimental class increased markedly from the pre-test to the post-test, while the control class showed only a slight increase. These descriptive results indicate that students who used Edbot.ai tended to achieve higher learning gains than those who learned through conventional written methods of diagnostic assessment. These descriptive results indicate that students who used Edbot.ai tended to achieve higher learning gains than those who learned through conventional written methods of diagnostic assessment.

Before conducting inferential statistical analyses using the Independent Samples t-test and N-Gain score analysis, statistical assumption tests were performed. These assumption tests included the normality test and the homogeneity of variance test, which are required to ensure that the data met the assumptions for applying parametric statistical procedures. Normality test is required to determine whether the data is normally distributed (Yang & Berdine, 2021). In this study, the normality test was conducted using the Shapiro–Wilk test with the assistance of SPSS. The criterion for the normality test was that if the significance value (Sig.) was greater than 0.05, the data were considered to be normally distributed. The results of the Shapiro–Wilk normality test for the pre-test and post-test scores of both the control and experimental classes are presented in Table 3. The findings indicate that all significance values obtained from the normality test were greater than 0.05 (Sig. > 0.05), suggesting that the data from both groups were normally distributed. Therefore, the assumption of normality was fulfilled, and the data were suitable for further analysis using parametric statistical tests.

Table 3. Shapiro-Wilk normality tests result

Data	Group	Shapiro-Wilk		
		Statistic	df	Sig.
Pre-test	Control Class	.983	30	.082
	Experimental Class	.955	30	.235
Post-test	Control Class	.958	30	.275
	Experimental Class	.953	30	.206

The next statistical assumption test conducted was the homogeneity of variance test. The homogeneity test aims to determine whether the variances of the two groups being compared are equal (Pallant, 2007). In this study, the homogeneity of variance was examined using Levene's Test. The criterion for the homogeneity test was that if the significance value (Sig.) was greater than 0.05, the data variances were considered homogeneous. The results of the homogeneity test for the pre-test and post-test scores of the control and experimental classes are presented in Table 4.

Table 4. Levene homogeneity test result

Data	Lavene Statistic	df1	df2	Sig.	Explanation
Pre-test	9.799	1	58	0.037	Not Homogenous
Post-test	23.769	1	58	0.000	Not Homogenous

The findings show that the significance values obtained from Levene's Test were less than 0.05 (Sig. < 0.05), indicating that the variances between the two groups were not homogeneous. Even though the data did not meet the homogeneity assumption, hypothesis testing can be validly conducted using independent sample t-test. Statistic solution for non-homogenous data is to refer to the t value in the "Equal variances not assumed" row (Welch's t-test), which is automatically provided by SPSS to accommodate the inequality of variances.

After the analytical assumption tests were fulfilled, the next step was hypothesis testing using the Independent Samples t-test and N-Gain score analysis. The Independent Samples t-test was conducted to determine whether there was a significant difference in students' learning outcomes between the control class and the experimental class on the post-test. The decision-making criterion was that if the significance value (Sig. 2-tailed) was less than 0.05, the difference in learning outcomes between the two groups was considered statistically significant. A summary of the t-test results is presented in Table 5.

Table 5. Independent sample t-test on post-test score result

Data	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
Post-test	-5,836	58	0.00	-21,50	3,684

Based on Table 5, the significance value (Sig. 2-tailed) obtained was 0.000, which is lower than 0.05. This result indicates that there was a statistically significant difference in English achievement between students who used Edbot.ai and those who were taught using conventional methods. Furthermore, the mean difference of –21.500 (the negative value resulted from SPSS calculating Group 1 minus Group 2) shows that the mean score of the experimental class was 21.50 points higher than that of the control class.

To further examine the effectiveness of Edbot.ai in improving students' English learning outcomes, an N-Gain score analysis was conducted. The N-Gain calculation compares the difference between post-test and pre-test scores with the maximum possible score that could be achieved. The summary of the mean N-Gain scores for both groups is presented in Table 6.

Table 6. N-Gain score mean comparison result

Group	N-Gain Mean	Interpreted Category
Control Class	0.04	Low
Experimental Class	0.53	Moderate

As shown in the Table 6, a clear difference in learning effectiveness can be observed between the two groups. The control class obtained a mean N-Gain score of 0.04, which falls into the low category, indicating only minimal improvement in students' English learning outcomes. In contrast, the experimental class achieved a mean N-Gain score of 0.53, which is categorized as moderate, showing a substantially higher level of learning improvement.

Furthermore, an Independent Samples t-test was conducted on the N-Gain data to examine whether the difference in learning improvement between the control and experimental classes was statistically significant. The results of the t-test are presented in Table 7.

Table 7. N-Gain Score's t-test Result

Data	t	df	Sig. (2-tailed)	Result
N-gain score	-6.105	39.40*	0	Significant

**Degrees of freedom (df) were adjusted using the "Equal variances not assumed" option due to unequal variances in the N-Gain data.*

In the table, the significance value (Sig. 2-tailed) was 0.000, which is lower than the significance level of 0.05. Therefore, it can be concluded that there was a statistically significant difference in learning improvement between the two groups. Since the assumption of homogeneity of variance was not met, the

analysis referred to the results in the "Equal variances not assumed" row, with the degrees of freedom adjusted accordingly. Overall, these findings indicate that the use of Edbot.ai was significantly more effective in improving students' English achievement than the conventional diagnostic assessment method. The learning improvement observed in the experimental class (N-Gain = 0.53) was substantially higher than that of the control class (N-Gain = 0.04).

Students' perception of Edbot.ai as a Diagnostic Assessment Tool

The qualitative data in this study were collected through an open-ended questionnaire administered via Google Forms consisting of eight questions given to 30 students of experimental class and were supported by classroom observations. This qualitative approach was employed to gain deeper insights into students' perceptions, experiences, and emotional responses toward the use of Edbot.ai as a diagnostic assessment tool in English learning. The questionnaire allowed students to freely express their opinions, while classroom observation provided additional contextual data regarding students' engagement and behavior during the learning process. The findings revealed several dominant themes related to students' engagement, perceptions of assessment, and learning support provided by Edbot.ai. Overall, most students expressed positive emotional responses, describing the use of Edbot.ai as fun, enjoyable, and engaging. These responses indicate that the integration of AI-based diagnostic assessment created a more interactive and motivating learning atmosphere compared to conventional written tests. All students' responses were translated into English to ensure clarity and consistency in data presentation. Student 14 said:

"Using Edbot.ai makes studying more fun and not boring."

Student 21 added:

> *"I think it is exciting because it is one of great AI applications that could ease us in learning English, especially because it has a lot of exercises and it could assess our speaking skill. "*

These statements suggest that the gamified and interactive nature of Edbot.ai contributed positively to students' motivation and engagement.

Despite the generally positive perceptions, some students acknowledged that the diagnostic tasks presented by Edbot.ai were moderately challenging. However, this level of difficulty was still perceived as appropriate and manageable. As Student 20 said:

"Some of the questions were easy but others were quite hard for me."

In line with Student 20's statement, Student 21 said:

"It was quite difficult but I still managed to answer it well."

This finding indicates that Edbot.ai was able to provide tasks that balanced challenge and accessibility, supporting students' learning without causing excessive frustration. This suggests that Edbot.ai successfully designed assessment items that challenged students cognitively while still remaining within

their zone of proximal development. Such balance is important in diagnostic assessment, as overly difficult tasks may discourage learners, while overly easy tasks may fail to accurately capture students' actual proficiency levels.

In term of assessment accuracy, most students stated that the diagnostic test results generated by Edbot.ai were consistent with their actual English abilities. Students felt that the platform accurately reflected their strengths and weaknesses across different language skills, particularly in grammar, vocabulary, listening, and speaking. This perception indicates that Edbot.ai functioned effectively as a diagnostic tool that helped learners gain clearer self-awareness of their current proficiency levels. It is in accordance to Student 6's statement:

> *"In my opinion, the assessment results from edbot.ai match the skills that I think I need to improve. It showed that my speaking skill is in low level."*

Student 22 also said:

> *"Edbot.ai is very helpful because I can find out which English skills need further improvement, in my case are listening and grammar."*

Beyond accuracy, students also expressed a strong preference for Edbot.ai over traditional written diagnostic tests. Many students described Edbot.ai-based assessment as more interactive, less tiring, and less time-consuming. Instead of perceiving the assessment as a formal and stressful test, students felt that completing tasks on Edbot.ai was similar to participating in a quiz or game. This shift in perception contributed to a more relaxed learning atmosphere. As Student 11 stated:

"Edbot.ai is great and fun and it is more practical than usual written test."

Student 18 added:

"*Edbot.ai is less-tiring because we only have to click the available options or respond orally through speaking, unlike traditional written tests*."

Furthermore, students perceived the course features of Edbot.ai as highly beneficial in supporting their learning process. The availability of immediate feedback, clear explanations, abundant practice opportunities, and flexible course selection allowed students to take greater control of their learning. Students appreciated the opportunity to choose courses based on the specific skills they needed to improve, which supported more personalized and self-directed learning. As Student 2 noted:

> *"The course feature is very useful since it gives us various material that we can choose by ourselves."*

In line with Student 2 statement, Student 17 added:

"Edbot.ai provides complete course features that allow students to focus on the specific language skills they want to develop. The direct feedback makes it easier to understand where our weaknesses directly."

Student 21 also said:

"Students are able to practice through a wide variety of questions with different levels and formats."

Classroom observations further reinforced these findings. During the learning process, students appeared more actively engaged, frequently asked questions, and interacted enthusiastically with the platform. Importantly, many students expressed that they did not feel burdened while completing the diagnostic tasks, as the activities felt more like playing than being tested. This observation aligns with students' questionnaire response of Student 12:

"If an assessment is conducted digitally, it is more fun because it is just like doing a quiz."

In accordance with Student 12, Student 19 also stated:

"It is really fun that I did not feel like I was studying and it just feels like I was playing a game."

Overall, the qualitative findings indicate that students responded positively to the use of Edbot.ai as a diagnostic assessment tool. Students perceived Edbot.ai as engaging, less intimidating, and more supportive of their learning needs compared to traditional written diagnostic tests. The platform not only helped students accurately recognize their English proficiency and areas for improvement, but also fostered a more relaxed and enjoyable learning atmosphere. The interactive features, immediate feedback, and flexible course selection encouraged active participation and reduced students' anxiety during assessment. These positive perceptions and learning experiences provide strong qualitative support for the quantitative findings, suggesting that the effectiveness of Edbot.ai in improving students' English achievement is closely related to increased engagement, motivation, and learner autonomy.

DISCUSSION

The use of Edbot.ai as a diagnostic assessment tool showed a positive contribution to students' English learning achievements as well as to their perceptions of the assessment process. The findings suggest that Edbot.ai supported students not only in identifying their learning progress quantitatively, but also in experiencing diagnostic assessment in a more engaging and meaningful way. This indicates that Edbot.ai functioned effectively as a diagnostic assessment tool that addressed both cognitive learning outcomes and students' affective responses.

To further understand this contribution, it is important to situate the findings within the broader discussion of AI-based diagnostic assessment in English language learning. Artificial intelligence–based assessment has been increasingly recognized as an effective approach for conducting diagnostic assessment in English language learning. Previous studies have shown that AI can support

diagnostic assessment by offering systematic and data-driven analysis of learners' language performance (Koraishi, 2023; Khan et al., 2021; Chen et al., 2020; Braiki et al., 2020). Diagnostic assessment is essential because it enables teachers to identify students' learning needs and design appropriate instructional interventions to address potential learning challenges (Adeoye et al., 2025). Furthermore, a well-implemented diagnostic assessment helps reveal gaps in students' language proficiency, allowing instruction to be more focused and targeted (Farisyah, 2025). In the context of this study, the AI-based diagnostic assessment was used to identify students' English abilities across four English skills: grammar, speaking, listening, and reading. With regard to grammar, AI systems are capable of detecting grammatical errors and identifying recurring patterns of learners' mistakes (Barata et al., 2025). For speaking skills, AI-assisted assessment facilitates diagnostic evaluation through speaking practice and automated analysis of pronunciation, fluency, and accuracy (Maulidyah, 2024; Taj, 2025). In the present study, the listening tasks embedded in Edbot.ai also included dictation activities, which required students to transcribe spoken input. Although previous studies mainly emphasize AI-assisted listening assessment for comprehension purposes (Dorgham, 2025), this study indicates that such tasks may simultaneously support spelling accuracy and basic writing skills. In addition, AI-based assessment supports reading comprehension by evaluating learners' ability to understand texts, identify main ideas, and make inferences (Chea & Xiao, 2024). Overall, the use of AI-based diagnostic assessment provides comprehensive information that supports teachers in planning subsequent instruction based on students' actual learning needs.

Beyond its diagnostic function, AI-based assessment offers a significant advantage through automated and immediate feedback. Previous studies indicate that AI systems are capable of automatically correcting and grading students' responses while providing feedback that highlights errors in learners' language performance (Maulidyah, 2024; Barata et al., 2025). This type of feedback supports formative assessment by enabling real-time monitoring of learners' performance during the learning process. From an instructional perspective, automated feedback allows teachers to identify students' difficulties without delay and respond with appropriate instructional adjustments. Such technology is particularly beneficial in classrooms with large numbers of students, where providing individual and timely feedback can be challenging (González-Calatayud, 2021). In addition, consistent feedback delivery ensures that assessment criteria are applied objectively across learners. This systematic feedback mechanism not only reduces teachers' workload but also enhances the efficiency of diagnostic assessment. Therefore, automated feedback functions as a

crucial component that strengthens the role of AI-based diagnostic assessment in English learning.

Taken together, these findings highlights that the use of AI-based diagnostic assessment can contribute positively to students' English achievement by supporting identification of learning needs and timely instructional responses. Beyond its impact on learning outcomes, the implementation of AI-based assessment was also well received by students, as reflected in their positive perceptions and learning experiences. This positive reception suggests that the effectiveness of AI-based diagnostic assessment is not only technical but also supported by students' engagement.

Students' positive perceptions were closely related to the learning atmosphere created through AI-based diagnostic assessment. AI-based diagnostic assessment was perceived as creating a more interactive learning environment and fostering a positive learning atmosphere for students. The integration of AI into assessment activities encouraged a more creative and engaging classroom environment, which has been shown to increase students' motivation in learning English (Maulidyah, 2024). Compared to traditional written assessments, AI-based assessment reduced the formality of testing and transformed assessment into a more learner-friendly activity. This supportive atmosphere helped students feel more relaxed during the learning process and minimized assessment-related anxiety (Dorgham, 2025). A positive learning atmosphere is essential in promoting students' willingness to participate and engage actively in learning tasks. When students perceive assessment as part of learning rather than a stressful evaluation, their emotional comfort increases. Consequently, AI-based diagnostic assessment contributes to a learning atmosphere that supports sustained engagement and positive learning experiences.

Beside engagement and emotional comfort, students also perceived the level of difficulty in AI-based diagnostic tasks as appropriate. AI-based diagnostic assessment, including the use of Edbot.ai, was perceived as providing tasks that were challenging yet still manageable for students. The adaptive nature of AI allows assessment tasks to vary in complexity, enabling students to engage with different levels of difficulty without feeling overwhelmed. This balance between challenge and accessibility is crucial in diagnostic assessment, as it encourages students to remain engaged while still revealing their actual learning needs. In line with this, Haenen et al. (2025) emphasize that the level of challenge, particularly in terms of task complexity and novelty, needs to be carefully managed to ensure that students feel both comfortable and motivated. Tasks that are too simple may reduce engagement, whereas tasks that are overly difficult can result in confusion and decreased motivation. In this context, AI-based assessment offers a supportive environment where students can attempt challenging tasks while still maintaining

confidence. Consequently, such an approach supports gradual improvement across language skills without causing excessive cognitive burden.

As students interacted continuously with diagnostic tasks and feedback, AI-based assessment also supported the development of learners' self-awareness. Edbot.ai practices function effectively as a diagnostic assessment by helping students develop greater awareness of their own language abilities. Through AI-based diagnostic assessment, students are not only informed of their performance but are also guided to recognize their strengths and areas that require improvement. AI systems are effective in delivering individualized feedback and allowing learners to adjust their learning pace based on their needs and preferred learning styles (Taj et al., 2025). This diagnostic feedback encourages learners to reflect on their learning progress and take a more active role in improving their language skills. Such reflective engagement is comparable to the role of peer assessment, which has been shown to promote students' proactivity in developing their language competence (Prastikawati & Adeoye, 2024). By continuously interacting with diagnostic results, students become more conscious of their learning gaps and progress over time. Moreover, personalized learning experiences facilitated through diagnostic assessment have been found to contribute significantly to better knowledge retention (Adeoye et al., 2025). Therefore, Edbot.ai not only supports skill diagnosis but also fosters students' self-awareness as a foundation for sustainable language development.

Another aspect highlighted by students was the practicality of AI-based diagnostic assessment compared to traditional testing methods. The use of Edbot.ai as a diagnostic assessment was perceived as more practical, efficient, and manageable for students. AI-based assessment allows learners to engage in more experiential learning activities through interactive digital tasks, which supports practical learning experiences when integrated with educational technology (Chen et al., 2020). This practical orientation reduces the need for lengthy written responses, enabling students to respond through selected options or oral input, thereby minimizing physical and cognitive fatigue. Several AI-based applications have also been reported to offer practical assessment formats that simulate real learning situations and promote active learner involvement (Amin, 2023). As a result, students are able to complete diagnostic tasks within a shorter time while maintaining engagement. The flexible and user-friendly digital interface further supports smooth classroom implementation without disrupting instructional time. Therefore, AI-based diagnostic assessment offers a more practical and experience-oriented alternative to conventional paper-based assessments.

While the practical and experience-oriented nature of Edbot.ai supports efficient task completion, its effectiveness is further strengthened by the feedback mechanism that guides students during and after the assessment process. Students

perceived the feedback mechanism embedded in Edbot.ai played a crucial role in supporting learning regulation and self-improvement. AI algorithms are able to analyze students' responses to questions and tasks and generate instant feedback based on their performance, allowing learners to immediately recognize their strengths and weaknesses (Amin, 2023). This process helps students identify specific areas that require improvement and encourages them to learn from their mistakes and make necessary adjustments. Rather than functioning solely as an evaluative tool, the feedback was perceived as instructional, guiding students toward better understanding of language concepts. Timely feedback also enabled students to monitor their own progress and take greater responsibility for their learning. As a result, students experienced a more reflective learning process in which feedback supported continuous improvement. Overall, AI-generated feedback was viewed as an essential component that enhanced students' awareness, autonomy, and engagement in English learning.

Concluding all perceptions and discussion, the implementation of AI-based diagnostic assessment demonstrates a meaningful contribution to English learning by supporting improved learning outcomes while also being positively perceived by students, as it creates an interactive learning atmosphere, provides appropriate levels of challenge, offers practical and efficient assessment procedures, and facilitates timely feedback that helps learners develop greater awareness of their language strengths and weaknesses.

CONCLUSION

The use of Edbot.ai as an AI-based diagnostic assessment contributes positively to English language learning by supporting the identification of students' learning needs, enhancing learning experiences, and fostering learner engagement and self-awareness. The integration of diagnostic features, automated feedback, and interactive tasks highlights the potential of AI-based assessment to support both learning outcomes and students' perceptions. The findings imply that Edbot.ai can serve as a complementary assessment tool that assists teachers in monitoring students' progress more efficiently and designing targeted instructional interventions based on diagnostic data. However, this study has several limitations, as it examined English skills in a general and integrated manner without focusing on a specific language skill, involved a relatively small number of participants, and was conducted within a short implementation period. These limitations may affect the depth and generalizability of the findings. Therefore, future research is recommended to investigate the long-term impact of AI-based diagnostic assessment involving a larger number of participants, investigate the use of Edbot.ai for specific English skills such as reading,

grammar, listening, or speaking, and explore its implementation in combination with different educational contexts and learning methods.

REFERENCES

Adeoye, M. A., Prastikawati, E. F., & Curle, S. (2025). Examining factors influencing language acquisition success in Indonesian multilingual classrooms. *Journal of Languages and Language Teaching*, *13*(4), 1743–1757.

Adeoye, M. A., Prastikawati, E. F., & Riwayatiningsih, R. (2025). Beyond traditional testing: Exploring the efficacy of adaptive assessments in shaping future learners. *Journal of Nonformal Education*, *11*(1), 153–169.

Aditama, M. G., Sugiharto, P. A., Istiqomah, L., & Hisyam, F. N. (2023). Integrating multiple intelligence tests into diagnostic assessment in ELT. *International Social Sciences and Humanities*, *2*(2), 358–363.

Al Braiki, B., Harous, S., Zaki, N., & Alnajjar, F. (2020). Artificial intelligence in education and assessment methods. *Bulletin of Electrical Engineering and Informatics*, *9*(5), 1998–2007. https://doi.org/10.11591/eei.v9i5.2470

Amin, M. Y. M. (2023). AI and ChatGPT in language teaching: Enhancing EFL classroom support and transforming assessment techniques. *International Journal of Higher Education Pedagogies*, *4*(4), 1–15.

Barata, A., Yuliana, Y. G. S., & Regina, R. (2025). Artificial intelligence chatbot application for improving teachers' spoken grammar. *Edukasi: Jurnal Pendidikan*, *23*(1), 163–178.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Chea, P., & Xiao, Y. (2024). Artificial intelligence in higher education: The power and risks of AI-assisted tools on academic English reading skills. *Journal of General Education and Humanities*, *3*(3), 287–306.

Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access*, *8*, 75264–75278. https://doi.org/10.1109/ACCESS.2020.2988510

Chen, Y. (2025). Transforming English language education with AI-driven deep learning models for scalable, adaptive, and inclusive assessment. *International Journal of Information and Communication Technology*, *26*(12), 15–31.

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.

Dorgham, R. (2025). Using artificial intelligence tools for developing EFL students' listening skills and reducing their listening anxiety. *Educational and Psychological Studies*, *40*(140, Part 1), 513–577.

Fang, Y., Roscoe, R. D., & McNamara, D. S. (2023). Artificial intelligence-based assessment in education. In *Handbook of artificial intelligence in education* (pp. 485–504). Edward Elgar Publishing.

Farhady, H., & Selcuk, M. (2022). Classroom-based diagnostic assessment practices of EFL instructors. *Iranian Journal of Language Teaching Research, 10*(2), 77–94.

Farisyah, U., Sulaimon, J. T., & Sodiq, J. (2025). Dynamic assessment strategies for enhancing reading comprehension and student motivation. *Educalingua, 3*(1), 1–12.

González-Calatayud, V., Prendes-Espinosa, P., & Roig-Vila, R. (2021). Artificial intelligence for student assessment: A systematic review. *Applied Sciences, 11*(12), 5467. https://doi.org/10.3390/app11125467

Haenen, J., Vink, S., Sjoer, E., & Admiraal, W. (2025). Motivating students with optimal challenge in an open learning environment. *Social Sciences & Humanities Open, 11*, 101577. https://doi.org/10.1016/j.ssaho.2025.101577

Jimola, F. E., & Ofodu, G. O. (2019). ESL teachers and diagnostic assessment: Perceptions and practices. *ELOPE: English Language Overseas Perspectives and Enquiries, 16*(2), 33–48.

Kazemi, N., & Tavassoli, K. (2020). The comparative effect of dynamic vs. diagnostic assessment on EFL learners' speaking ability. *Research in English Language Pedagogy, 8*(2), 223–240.

Khan, I., Ahmad, A. R., Jabeur, N., & Mahdi, M. N. (2021). An artificial intelligence approach to monitor student performance and devise preventive measures. *Smart Learning Environments, 8*(1), 17. https://doi.org/10.1186/s40561-021-00161-y

Koraishi, O. (2023). Teaching English in the age of AI: Embracing ChatGPT to optimize EFL materials and assessment. *Language Education and Technology, 3*(1), 1–10.

Lee, D., Kim, H. H., & Sung, S. H. (2023). Development research on an AI English learning support system to facilitate learner-generated-context-based learning. *Educational Technology Research and Development, 71*(2), 629–666. https://doi.org/10.1007/s11423-022-10172-2

Luo, J., Zheng, C., Yin, J., & Teo, H. H. (2025). Design and assessment of AI-based learning tools in higher education: A systematic review. *International Journal of Educational Technology in Higher Education, 22*(1), 42.

Maulidyah, D. S. (2025). Exploring teachers' perceptions of AI-powered classroom-based assessment tools in ESP class. *Indonesian Journal of Foreign Language Studies, 2*(1), 28–36.

Mohammadi, M. (2024). Language teachers' assessment literacy in AI-aided adaptive learning environments. *Journal of Research in Applied Linguistics, 15*(2), 73–88.

Nuraini, F. I., Sodiq, J., & Prastikawati, E. F. (2025). Edmodo as technology-based formative assessment: How it enhances reading comprehension. *International Journal of Research in Education, 5*(1), 81–92.

Pallant, J. (2007). *SPSS survival manual: A step-by-step guide to data analysis using SPSS* (3rd ed.). McGraw-Hill Education.

Prastikawati, E. F., & Adeoye, M. A. (2024). Enhancing learning: Peer assessment's influence on English as a foreign language education in Indonesia. *Mimbar Ilmu, 29*(2), 246–253.

Rahayu, W., Prastikawati, E. F., Wiyaka, W., & Lestari, M. Y. W. (2022). Fostering students' reading comprehension through dynamic assessment. *Language Circle: Journal of Language and Literature, 17*(1), 205–214.

Razali, N. M., & Wah, Y. B. (2011). Power comparisons of Shapiro–Wilk, Kolmogorov–Smirnov, Lilliefors, and Anderson–Darling tests. *Journal of Statistical Modeling and Analytics, 2*(1), 21–33.

Setiyowati, R., & Ardaniah, V. (2025). A systematic review of AI-based and teacher-based writing assessment. In *Proceedings of the English Language and Literature International Conference (ELLiC)* (Vol. 8, pp. 380–394).

Swiecki, Z., Khosravi, H., Chen, G., Martinez-Maldonado, R., Lodge, J. M., Milligan, S., & Gašević, D. (2022). Assessment in the age of artificial intelligence. *Computers and Education: Artificial Intelligence, 3*, 100075. https://doi.org/10.1016/j.caeai.2022.100075

Taj, A., Khan, S., Siddiqui, H. A., Abbas, Q., & Nawaz, A. (2025). Artificial intelligence in addressing vocabulary learning challenges: Strategies for English language learners. *The Critical Review of Social Sciences Studies, 3*(1), 3321–3342.

Tosuncuoglu, I. (2018). Importance of assessment in ELT. *Journal of Education and Training Studies, 6*(9), 163–167. https://doi.org/10.11114/jets.v6i9.3443

Wiyaka, W., Silitonga, L. M., Sunardi, S., & Pramudi, Y. T. C. (2025). Leveraging diagnostic assessment aligned to differentiated learning: Voices of English language teachers. *KnE Social Sciences*.

Yang, S., & Berdine, G. (2021). Normality tests. *The Southwest Respiratory and Critical Care Chronicles, 9*(37), 87–90.

Yang, Y. (2025). AI-supported L2 vocabulary acquisition: A systematic review from 2015 to 2023. *Education and Information Technologies*. Advance online publication.