FRASA: English Education and Literature Journal
Vol. 6 No. 2, September 2025

ISSN: 2807-8195
Available online at: https://ojs.udb.ac.id/
DOI: https://doi.org/10.47701/frasa.v6i2.5037

ARTIFICIAL INTELLIGENCE (AI)-BASED ASSESSMENT: A TRANSFORMATIVE ENGLISH EDUCATION CURRICULUM

[1]Benni Ichsanda Rahman Hz, [1]Ammamiarihta

[1]Universitas Islam Negeri Sumatera Utara, Indonesia

***Corresponding Author**
Email: benni.ichsanda@uinsu.ac.id

Received:	Revised:	Accepted:	Published:
16 June 2025	7 July 2025	15 August 2025	4 September 2025

How to cite (APA 7th style): Rahman Hz, B. I., & Ammamiarihta, A. (2025). Artificial intelligence (AI)-based assessment: A transformative English education curriculum. *Frasa: English Education and Literature Journal, 6*(2), 289–303. https://doi.org/10.47701/frasa.v6i2.5037

Abstract

Artificial Intelligence (AI) has been developing at a fast pace and is increasingly embedded in everyday human activities. In the realm of education AI serves as a transformative force, reshaping both teaching practices and learning processes. A key contribution of AI in this context lies in its capacity to deliver individualized and adaptive learning experiences. This study aims to develop an Artificial Intelligence (AI)-based English learning assessment system in the English Education program at the State Islamic University of North Sumatra, Medan. The study also seeks to measure the extent to which the AI-based assessment curriculum has effectively achieved the program's vision, mission, objectives, and graduate profile. A qualitative case study within purposive sampling approach was employed to select participants who had relevant experience and knowledge in English teaching and learning. The data collection methods included Focused Group Discussions (FGDs) with 20 lecturers and 100 students, as well as in-depth interviews with lecturers. The data from questionnaire were analyzed to identify patterns, trends, and differences in the responses of the respondents. The findings of this study indicate that the use of AI in assessing English proficiency among students has proven to be effective and efficient, not only in individual assessment but also in supporting the evaluation of academic program curricula. However, the findings also highlight the need for adequate technological readiness and training for both lecturer and student to optimize AI utilization.

Keywords: *Artificial Intelligence, Assessment, Curriculum, English*

INTRODUCTION

Artificial Intelligence (AI) is rapidly evolving, becoming an integral part of human life today (Chiu et al., 2022; Schepman & Rodway, 2023; Wang et al., 2023). Including in the field of education (Nazari et al., 2021), AI plays a crucial role in transforming the way people teach and learn. One of AI's significant contributions to education is its ability to provide

FRASA: English Education and Literature Journal
Vol. 6 No. 2, September 2025
ISSN: 2807-8195
Available online at: https://ojs.udb.ac.id/
DOI: https://doi.org/10.47701/frasa.v6i2.5037

personalized and adaptive learning experiences (Kabudi et al., 2021; Delgado et al., 2020; Zawacki-Richter et al., 2019). Through big data analysis and machine learning, AI can understand students' learning patterns, identify individual needs, and present tailored learning content. This enables each student to learn at their own pace and style, enhancing the effectiveness of learning. Additionally, AI is used in the development of interactive and responsive e-learning platforms (Tang et al., 2023).

Moreover, AI systems can provide instant feedback, adjust the curriculum based on each student's strengths and weaknesses, and offer challenges appropriate to their level of ability. In assessment and evaluation, AI delivers speed and accuracy unattainable with manual methods (Balakumar et al., 2021; Xu et al., 2021). Automated systems for test evaluation, student performance analysis, and instant feedback accelerate the learning process and help teachers provide additional support to students in need. AI also plays a role in supporting the development of more dynamic and relevant curricula. By analyzing global educational trends and industry needs, AI can assist in designing curricula that integrate real-world skills.

In the field of English Language Education, AI can be used and developed to meet the needs of proficiency test development. The use of English proficiency tests has been a significant part of the education system for centuries, and their development and application have significantly impacted the formation of language learning and proficiency (Frost et al., 2020; Owen et al., 2021). The development of language proficiency tests is closely related to the growth of globalization and the need for a common language of communication in international contexts, which over time has led to an increasing demand for language proficiency testing (Bahman & Adrian, 2022).

English proficiency tests are often used to measure the general English language skills of users, whether they are students, workers, scholarship seekers, or immigrants. However, English proficiency tests are still rarely used separately to assess the capabilities of students in the field of English Language Education. These students are future English teachers who will be teaching in various educational institutions. Therefore, the English Language Education curriculum is specifically designed to equip prospective teachers with the skills necessary to teach English effectively. In assessing these skills, English proficiency tests for prospective teachers play an essential role. These tests not only evaluate the linguistic abilities of future teachers but also focus on specific aspects, such as effective communication, pedagogical skills, grammar knowledge, listening ability, and the assessment of students' needs.

The measurement of English proficiency among students is still conducted manually by lecturers and program administrators. The lack of automated tools leads to subjective assessments and a lack of consistency. The implementation of a structured system with automated guidelines is expected to improve the validity and reliability of assessments, supporting the enhancement of teaching and learning quality in English Language Education.

This study aims to develop an Artificial Intelligence (AI)-based English learning assessment system in the English Education program at the State Islamic University of North Sumatra, Medan. The study also seeks to measure the extent to which the AI-based assessment curriculum has effectively achieved the program's vision, mission, objectives, and graduate profile.

FRASA: English Education and Literature Journal
Vol. 6 No. 2, September 2025

ISSN: 2807-8195
Available online at: https://ojs.udb.ac.id/
DOI: https://doi.org/10.47701/frasa.v6i2.5037

Previous research shows that English proficiency tests are accurate and valid measures of English language proficiency (e.g., (Kyle et al., 2016; Zahedi & Shamsaee, 2012; Zahedkazemi, 2015). Specifically, Kyle et al., (2016) found that certain linguistic features, such as lexical diversity, syntactic complexity, and discourse coherence, positively correlate with overall speaking test scores, indicating that speaking tests like TOEFL iBT have strong construct validity. Similarly, Zahedi & Shamsaee (2012) demonstrated a highly significant correlation between speaking scores on the TOEFL iBT and IELTS, suggesting that both tests measure the same oral proficiency construct. Furthermore, Zahedkazemi (2015) emphasized the importance of ensuring that English proficiency tests are valid and reliable measures for test-takers from diverse linguistic and cultural backgrounds.

In the context of AI development in learning, Evanini et al. (2017) researched the use of automatic speech assessment for K-12 English Language Proficiency (ELP), focusing on assessment construction, task design, assessment reporting, AI model development, and test delivery. They recommended that certain countries adopt tools for evaluating proficiency test users. In another study, El Shazly (2021) explored using an AI chatbot to manage foreign language anxiety among 48 English majors in Egypt over an eight-week period. The results showed a slight increase in anxiety, but the AI chatbot showed promise in improving linguistic outcomes in language learning. Additionally, Gayed et al., (2022) developed an AI-based web application called "AI KAKU" to help English learners reduce cognitive barriers when producing written texts in English. Early results indicated that their tool could benefit English learners needing more structured writing assistance than traditional word processors. Furthermore, Huang & Flores (2018) evaluated the English Language Proficiency Assessment for the 21st Century (ELPA 21), a large-scale English proficiency test used in eight states, which was designed to help English learners acquire the academic language skills necessary for college and career readiness.

Regarding AI-based English education curriculum development, Bin & Mandal (2019) explored the application of AI in teaching English at secondary schools using relevant theories from curriculum theory and literature analysis. They implemented AI-assisted English teaching systems to enhance the quality and effectiveness of English instruction. An et al., (2023) examined English teachers' perceptions, knowledge, and behavioural intentions in China regarding the use of AI to support English teaching and learning at the secondary level. Their findings highlight the complex relationships underlying teachers' attitudes toward using AI in education.

The first modern language proficiency tests were developed in the early 20th century, and their popularity grew rapidly as more people began traveling, working, and studying in countries where English is spoken (Fernandez, 2018; Fulcher, 2012). English proficiency tests play a crucial role in language learning and have been extensively studied in the fields of applied linguistics and language education. These tests are designed to assess users' language proficiency and are applied for various purposes, including university admissions, job applications, and immigration requirements.

Currently, the most popular English proficiency tests are TOEFL, administered by the English Testing System; IELTS, administered by the British Council; and PTE Academic,

FRASA: English Education and Literature Journal
Vol. 6 No. 2, September 2025
ISSN: 2807-8195
Available online at: https://ojs.udb.ac.id/
DOI: https://doi.org/10.47701/frasa.v6i2.5037

administered by Pearson Education. However, these popular tests are still conducted manually, meaning that their implementation, measurement, and assessment are also performed manually. Many educational institutions use these models to measure their students' English proficiency manually.

In practice, the measurement of English proficiency among students is still conducted manually by individual instructors and program administrators. This means assessments are based on the subjective evaluations of instructors and comprehensive exams at the end of the semester, without the support of automated tools or more detailed evaluation standards. These limitations can lead to subjectivity in assessments and a lack of consistency among evaluators. Therefore, implementing a more structured English proficiency measurement system, using automated tools or established assessment guidelines, can help improve the validity and reliability of evaluations, providing a more accurate picture of students' English proficiency. Efforts toward more standardized and objective measurements can support continuous improvement in the quality of teaching and learning in English education programs.

With the advancement of artificial intelligence (AI) technology, AI is crucial for assessing students' English proficiency (Koraishi, 2023; Rukiati et al., 2023). AI offers significant advantages in the evaluation process. First, the use of AI can facilitate smarter decision-making (Albiladi & Alshareef, 2019; Yang et al., 2011), enhancing objectivity and consistency in assessments by reducing the influence of human subjectivity. AI systems can be designed to provide consistent evaluations bsased on predetermined criteria, yielding more accurate and reliable results (Mikalef & Gupta, 2021). Second, AI can deliver assessments quickly and efficiently, enabling faster evaluations than traditional manual methods (Manongga et al., 2022). This not only reduces the workload for instructors and program administrators but also provides faster feedback to students, allowing them to quickly identify areas needing improvement.

Moreover, AI technology can be configured to present more interactive and engaging assessment tasks, creating a more dynamic learning experience for students. By leveraging AI, assessment tasks can be tailored to each student's ability level, allowing for a personalized and effective learning approach. Additionally, the use of AI in assessing English proficiency can help identify patterns and trends in students' overall performance. This information can be used to enhance curriculum design, adjust teaching strategies, and provide more targeted feedback to instructors for further improvements.

The use of AI in assessing students' English proficiency is also expected to align with the evaluation of the curriculum itself. With AI, curriculum achievements can be measured more quickly (Martínez-Comesaña et al., 2023; Swiecki et al., 2022), providing a more accurate picture of how well the objectives and competencies of the curriculum are being met. AI systems can generate real-time evaluation data (Muhammad Yahya et al., 2023), enabling program administrators to track student progress and identify strengths and weaknesses within the curriculum more efficiently (Martínez-Comesaña et al., 2023). Furthermore, integrating AI into English proficiency assessments can help identify trends in overall student performance, allowing for more informed decision-making regarding curriculum adjustments.

FRASA: English Education and Literature Journal
Vol. 6 No. 2, September 2025
ISSN: 2807-8195
Available online at: https://ojs.udb.ac.id/
DOI: https://doi.org/10.47701/frasa.v6i2.5037

Through in-depth data analysis, the use of AI allows for more precise adjustments to the teaching materials and methods employed in the curriculum (Chen et al., 2022). Additionally, utilizing AI in curriculum evaluation presents opportunities for developing a more adaptive and responsive curriculum. The data collected by AI can be used to design quicker and more accurate curriculum changes, accommodating the evolving needs of students and the changing demands in the field of English language professions.

Previous studies have explored the use of AI in general language proficiency testing (e.g., Evanini et al., 2017; Gayed et al., 2022) and its application in secondary education (Bin & Mandal, 2019; An et al., 2023). This study, on the other hand, is the initial attempt to study AI-based assessment in an English Education program for future teachers. It fills an important gap by looking at its function in not just evaluating individual competency but also aligning the curriculum and program goals. It also uses both questionnaire and FGD to give a complete picture of how well AI works, while also pointing out the problems with measuring subtle language skills that are important for future teachers. This gives new ideas for how to train teachers and use AI in higher education.

RESEARCH METHODOLOGY

This study adopted a qualitative case study method with the primary aim of measuring the effectiveness of AI-based English learning assessments. Qualitative case studies were used to explore a phenomenon in its natural setting, employing various data sources to ensure a comprehensive understanding of the issues while avoiding a singular perspective and considering multiple complexities (Baxter & Jack, 2008). A purposive sampling approach was employed to select participants who had relevant experience and knowledge in English teaching and learning. Data collection methods included Focused Group Discussions (FGDs) with 20 lecturers and 20 students, as well as in-depth interviews with lecturers. Additionally, questionnaires were distributed to 20 lecturers and 100 students to gain broader insights. The participation of lecturers and students in FGDs and interviews was expected to provide deep insights into their experiences with AI-based assessments, while the questionnaires offered quantitative data for further analysis. The questionnaires distributed to 20 lecturers and 100 students were developed based on a review of prior AI-in-education literature (e.g., Kabudi et al., 2021; Gayed et al., 2022) and customized to assess perceptions of AI-based assessments in English proficiency testing. Sample items included Likert-scale questions (e.g., "How effective is AI in evaluating English proficiency?") and open-ended queries (e.g., "Describe challenges you faced with AI assessments"), capturing both quantitative and qualitative information. Focus group discussions (FGDs) and interviews investigated more closely at issues including usability, accuracy, and ethics, using questions based on the results of the first questionnaire.

A total of 20 lecturers and 100 students participated as respondents in this study. The lecturers involved in the FGDs and interviews were expected to provide their perspectives as English instructors and direct users of AI-based assessments. Meanwhile, student participation in the questionnaires offered a view from the learner's side. The relatively large number of respondents was anticipated to provide a diversity of opinions and experiences, forming a basis for a comprehensive analysis of the effectiveness of AI-based assessments in English learning.

FRASA: English Education and Literature Journal
Vol. 6 No. 2, September 2025
ISSN: 2807-8195
Available online at: https://ojs.udb.ac.id/
DOI: https://doi.org/10.47701/frasa.v6i2.5037

The data analysis techniques for this research encompassed two complementary approaches. First, quantitative analysis was conducted through content analysis of the questionnaire responses from 20 lecturers and 100 students. This quantitative data was statistically analysed to identify patterns, trends, and differences in the responses of lecturers and students regarding the effectiveness of AI-based assessments in English learning. This analysis provided a more measurable understanding of the perceptions and experiences of the respondents, highlighting specific aspects that needed attention in the context of using AI technology in assessments.

Simultaneously, qualitative analysis was carried out through the interpretation of interview results with lecturers. These in-depth interviews provided profound insights into the views and experiences of lecturers regarding the implementation of AI-based assessments. Qualitative data from the interviews was analysed to identify key findings, patterns, and nuances that may not have been revealed through quantitative analysis. The integration of both approaches was expected to provide a holistic and in-depth picture of the effectiveness of AI-based assessments in English learning, combining quantitative and qualitative dimensions for a more comprehensive understanding.

RESULTS AND DISCUSSION

Results

In the first stage, the researcher distributes a questionnaire to respondents, with a total of 40 individuals providing responses, consisting of 36 students and 4 lecturers. The questionnaire, containing 20 items, is designed to gather data regarding their perceptions, experiences, and views on the use of AI in the implementation of the English Proficiency Test. The questionnaire utilizes a Likert scale with the following details:

Table 1. The descriptions of Likert Scale

Number	Descriptions
5	Strongly Agree
4	Agree
3	Neutral
2	Disagree
1	Strongly Disagree

With this questionnaire, the researcher aims to understand the extent of the effectiveness, efficiency, and security of using AI in the test, as well as to evaluate the potential and challenges faced by users. The data obtained from the respondents will be analysed to provide deeper insights and support further development in the implementation of AI in English proficiency testing. Based on the results of the questionnaire, the following data are obtained:

Table 2. The Results of The Questionnaire

No	Statement	Score				
		5	**4**	**3**	**2**	**1**
1	How effective do you think the use of AI is in the implementation of the English Proficiency Test/Assessment?	40%	50%	7,5%	0%	2,5%

FRASA: English Education and Literature Journal
Vol. 6 No. 2, September 2025
ISSN: 2807-8195
Available online at: https://ojs.udb.ac.id/
DOI: https://doi.org/10.47701/frasa.v6i2.5037

2	How easy do you think it is to use the AI system in the English Proficiency Test/Assessment?	45%	37,5%	15%	2,5%	0%
3	How accurate do you think the AI system is in assessing English language proficiency?	10%	52,5%	32,5%	5%	0%
4	How safe do you think the personal data collected by the AI system during the English proficiency test is?	7,5%	37,5%	42,5%	12,5%	0%
5	How relevant do you think the use of AI is in the English proficiency test/assessment today?	30%	50%	15%	5%	0%
6	How familiar are you with the use of AI in the English proficiency test?	35%	37,5%	17,5%	10%	0%
7	How often do you use or participate in English proficiency tests/assessments that utilize AI?	17,5 %	25%	47,5%	7,5%	2,5%
8	To what extent do you trust the accuracy of AI assessments in English proficiency tests?	10%	52,5%	35%	2,5%	0%
9	How relevant do you think the use of AI is in improving the quality of English proficiency tests/assessments?	32,5 %	62,5%	5%	0%	0%
10	How efficient do you think the use of AI is in English proficiency tests/assessments compared to conventional methods?	35%	37,5%	25%	2,5%	0%
11	How significant do you think the impact of using AI in English proficiency tests/assessments is on costs?	30%	32,5%	25%	12,5%	0%
12	How easy do you think it is to adapt to the use of AI in English proficiency tests/assessments?	35%	45%	17,5%	2,5%	0%
13	How significant is the contribution of AI in reducing the time taken for English proficiency tests/assessments?	57,5 %	30%	12,5%	0%	0%
14	How important do you think training and socialization on the use of AI are for users of English proficiency tests/assessments?	52,5 %	32,5%	15%	0%	0%
15	How concerned are you about the security of personal data in the use of AI in English proficiency tests/assessments?	10%	27,5%	47,5%	10%	5%
16	How hopeful are you about the development of AI in enhancing the quality of English proficiency tests in the future?	57,5 %	30%	12,5%	0%	0%
17	How confident are you that the use of AI in English proficiency tests can enhance the user experience?	50%	35%	10%	5%	0%
18	How strongly do you feel that AI can help address individual needs in English proficiency tests?	45%	42,5%	12,5%	0%	0%
19	How strongly do you feel that AI can help minimize bias in the assessment of English proficiency tests?	25%	37,5%	35%	2,5%	0%

FRASA: English Education and Literature Journal
Vol. 6 No. 2, September 2025
ISSN: 2807-8195
Available online at: https://ojs.udb.ac.id/
DOI: https://doi.org/10.47701/frasa.v6i2.5037

20	How strongly do you agree that the use of AI in English proficiency tests can enhance the efficiency of the education system?	40%	40%	20%	0%	0%

Advantages

Overall, based on the results of the questionnaire, the majority of respondents welcomed the presence of AI in English language assessment. In terms of effectiveness, most respondents agreed that AI is effective in conducting the English Proficiency Test. This is evident from the questionnaire results, where 40% of respondents strongly agreed that AI is effective, and 50% agreed with that statement. About 7.5% of respondents rated it neutral, while no respondents disagreed. Only 2.5% strongly disagreed. This data indicates that more than 90% of respondents believe that the use of AI in the English Proficiency Test has a high level of effectiveness, reflecting strong confidence in AI's ability to enhance the quality and accuracy of English tests.

Similarly, in terms of efficiency, the majority of respondents are confident that AI in assessments is very efficient in reducing the use of other resources. This is reflected in the questionnaire results, where 40% strongly agreed and another 40% agreed that AI is efficient. About 20% of respondents rated it neutral, while no respondents disagreed or strongly disagreed. This data shows that 80% of respondents believe that the use of AI in the English Proficiency Test can improve efficiency by reducing the need for additional resources, such as time and effort, reflecting strong confidence in AI's potential to enhance assessment efficiency.

In terms of ease of use, the majority of respondents also found AI to be very user-friendly in the English language assessment process. This is evident from the questionnaire results, where 45% of respondents strongly agreed that AI is easy to use, and 37.5% agreed. About 15% rated it neutral, while only 2.5% disagreed, and no respondents strongly disagreed. This data indicates that over 80% of respondents consider the use of AI in English language assessment to be easy and user-friendly, reflecting strong confidence in AI's ability to simplify the assessment process.

In contrast, regarding the accuracy of assessments using AI, respondents had varied opinions, though still leaning positively. Based on the question, "To what extent do you trust the accuracy of AI's assessment in the English Proficiency Test?" the questionnaire results showed that 10% of respondents strongly trust AI's accuracy, while 52.5% trust it. About 35% rated it neutral, and only 2.5% do not trust AI's accuracy. No respondents strongly distrust it. This data indicates that despite some variation in opinions, the majority still believe that AI can provide accurate assessment results in English proficiency tests.

In terms of data privacy security, respondents showed varying but significant levels of concern. Based on the question, "How concerned are you about the privacy of personal data in the use of AI in the English Proficiency Test?" the questionnaire results showed that 10% of respondents are very concerned, 27.5% are concerned, and 47.5% rated it neutral. About 10% are not concerned, and 5% are very unconcerned. This data indicates that while some respondents are not very worried, there is still a significant concern among respondents regarding the privacy of personal data in the use of AI for English assessments.

FRASA: English Education and Literature Journal
Vol. 6 No. 2, September 2025
ISSN: 2807-8195
Available online at: https://ojs.udb.ac.id/
DOI: https://doi.org/10.47701/frasa.v6i2.5037

Remarkable Shortcomings

However, in interviews and focused group discussions, respondents pointed out several shortcomings they experienced.

"To me, the drawback of AI in assessing someone's English proficiency is that AI can be easily distracted by external disturbances, which can lead to biased assessments. AI is also less effective in assessing context and nuance in a language because language is not just about the words spoken; it also has feelings and emotions that are difficult for AI to evaluate." - Respondent 1

"In my opinion, there are several drawbacks of AI in assessing English proficiency, namely, frequent errors in speech recognition when performing speaking skills, inaccuracies in the data results, and the AI on the smalltalk2.me website has a pause menu that is inefficient because users might use it to search for answers before continuing the recording. Thus, their actual ability does not show." - Respondent 2

"I believe AI is still lacking in assessing someone's language skills because each individual has a unique and character-driven language style. This limitation means AI cannot identify and understand such languages. Additionally, AI struggles to comprehend culturally contextual phrases or idioms, leading to a loss of meaning in the text when using AI." - Respondent 3

"A downside is that sometimes when using AI, there are issues, like with smalltalk2.me, where I couldn't figure out why the audio wouldn't play. I tried repeatedly, and the result was the same—it couldn't be played. I didn't know where the problem was, so I continued the test without filling that part, leading to an inaccurate score because I didn't complete that section. There were also several words I pronounced, but AI picked up different words, possibly due to background noise." - Respondent 4

"Sometimes, perhaps due to the conditions when, for example, we are practicing speaking with AI, if the environment is not conducive, whether due to noise or unstable internet connection, it can cause errors in the system, meaning what we said may be correct but is interpreted incorrectly by the system due to those factors." - Respondent 5

"In my view, the limitation of AI in assessing someone's English proficiency is primarily in pronunciation skills, as sometimes the words or sentences detected by AI do not match what we actually said." - Respondent 6

"AI has several shortcomings in assessing someone's English proficiency. AI often struggles to understand cultural contexts, idioms, or varying expressions in the language. Dependence on training data also poses a barrier, as AI may not be able to handle variations or errors not present in that data. AI cannot assess aspects of creativity and intuition in language use, which are often crucial in overall language proficiency assessments." - Respondent 7

"I believe AI has several limitations in assessing someone's English proficiency, first, Context Limitations; AI often struggles to understand broader contexts in conversations or texts. This can lead to errors in assessment, especially if the text contains deep meanings or cultural connotations. Second, Difficulty in Assessing

FRASA: English Education and Literature Journal
Vol. 6 No. 2, September 2025

ISSN: 2807-8195
Available online at: https://ojs.udb.ac.id/
DOI: https://doi.org/10.47701/frasa.v6i2.5037

Creativity; English proficiency encompasses creativity in language use, such as metaphor creation or humour, which is challenging for AI to measure. Third, sensitivity to technical errors; AI tends to overly emphasize grammar or spelling mistakes while neglecting other aspects of language proficiency that may be more important in certain contexts, such as fluency and clarity. But overall, despite some shortcomings, AI is very helpful as a tool for English assessment now and in the future." - Respondent 8

Based on interviews and focused group discussions, respondents identified several shortcomings in using AI to assess English proficiency. First, AI is easily disrupted by background noise or other external disturbances, leading to biased assessments. This indicates that AI still has limitations in handling less-than-ideal testing environments and may require further development to improve its resilience against external disruptions. Second, AI often struggles to assess context and emotion in language, which are important for fully understanding someone's language abilities. Language is not just words; it also encompasses emotional and contextual expressions that are challenging for AI to measure.

Additionally, respondents noted that AI frequently makes errors in speech recognition, particularly when assessing speaking abilities. These inaccuracies may arise from various factors, including AI's inability to understand unique and character-driven language styles. Another issue identified is the inefficiency of certain features on AI platforms, such as the pause menu on smalltalk2.me, which allows users to search for answers before continuing their recording, thus not reflecting their true capabilities.

AI struggles to understand cultural contexts

Furthermore, AI struggles to understand cultural contexts, idioms, and varied expressions in the language. Dependence on training data also becomes a barrier, as AI may not be able to handle variations or errors that are not present in that data. AI also cannot assess aspects of creativity and intuition in language use, which are often important in evaluating overall language proficiency. Additionally, sensitivity to technical errors, such as grammar or spelling mistakes, can obscure other important aspects of language proficiency, such as fluency and clarity. Nevertheless, the majority of respondents agreed that despite its shortcomings, AI remains a very useful tool in English language assessment today and in the future. Meanwhile, the lecturers present at the FGD expressed similar views.

"I haven't used AI; I still teach listening with audio, but I will delve deeper into the assessments available in AI." - Lecturer 1

"AI in ICT is very interesting; with AI, I can directly check and assess. AI is very beneficial in assessments." - Lecturer 2 (ICT Lecturer)

"I recommend using AI in writing. As future English teachers, we must be able to leverage AI in education." - Lecturer 3

At the same time, the lecturers present at the FGD expressed similar opinions regarding the use of AI in English language learning. Lecturer 1 stated that even though they haven't used AI in their teaching yet, they still use audio to teach listening skills. However, they plan to explore the assessment potential that AI can provide. Lecturer 2, who teaches in the ICT field, expressed that AI in ICT is very interesting and beneficial, as it allows for direct checking and assessment. They emphasized that AI is very helpful in the assessment process.

FRASA: English Education and Literature Journal
Vol. 6 No. 2, September 2025
ISSN: 2807-8195
Available online at: https://ojs.udb.ac.id/
DOI: https://doi.org/10.47701/frasa.v6i2.5037

Lecturer 3 recommended the use of AI in writing skills. They always suggest using AI in teaching writing and believe that as English teachers, it is essential to utilize AI technology in education. These opinions indicate that despite some concerns regarding AI's shortcomings, the lecturers see great potential in its use to improve efficiency and accuracy in English language assessment.

Discussion

Overall, based on the experiences of the respondents involved, AI-based assessment has proven to be effective and efficient in the process of learning and evaluating an individual's English proficiency. This aligns with the theory that in the field of education (Nazari et al., 2021), AI plays a significant role in transforming the way people teach and learn. One of AI's major contributions to education is its ability to provide personalized and adaptive learning experiences (Kabudi et al., 2021; Delgado et al., 2020; Zawacki-Richter et al., 2019). By utilizing big data analysis and machine learning, AI can understand students' learning patterns, identify individual needs, and deliver tailored learning content. This allows each student to learn according to their own level and style, thereby enhancing learning effectiveness. Furthermore, AI is also leveraged in the development of interactive and responsive e-learning platforms (Tang et al., 2023).

In the context of assessment and evaluation, AI offers significant advantages that cannot be achieved with traditional manual methods. With the ability to process data quickly and accurately, AI enables more efficient assessments, significantly reducing the time needed to evaluate student performance. For instance, AI can automatically analyze and assess written tasks or oral exams with consistent precision, reducing the likelihood of human error and subjective bias in manual grading (Balakumar et al., 2023; Xu et al., 2021). This technology is also capable of managing and evaluating large volumes of data quickly, making it ideal for large-scale assessments. Another advantage of AI in assessment is its ability to provide real-time feedback to students. This rapid feedback allows students to immediately recognize areas they need to improve, accelerating their learning process. Thus, AI not only improves efficiency and accuracy in assessment but also contributes to a more dynamic and responsive learning experience (Balakumar et al., 2023; Xu et al., 2021).

This statement also aligns with the view that, with the advancement of AI technology, its use becomes increasingly important in assessing the English proficiency of students (Koraishi, 2023; Rukiati et al., 2023). This is due to the various significant benefits AI offers in the evaluation process. First, the application of AI can lead to smarter decision-making (Albiladi & Alshareef, 2019; Yang et al., 2011) and enhance the objectivity and consistency of assessments by reducing the subjectivity of lecturers or human evaluators. AI systems can be programmed to provide consistent evaluations based on predetermined criteria, resulting in more accurate and reliable outcomes (Mikalef & Gupta, 2021).

Moreover, the findings of this study support the belief that using AI in the English proficiency assessment of students can be aligned with curriculum evaluation in academic programs. The presence of AI allows for faster measurement of curriculum achievement (Martínez-Comesaña et al., 2021; Swiecki et al., 2022), providing a more accurate picture of the extent to which the curriculum's goals and competencies have been achieved. AI systems have

FRASA: English Education and Literature Journal
Vol. 6 No. 2, September 2025
ISSN: 2807-8195
Available online at: https://ojs.udb.ac.id/
DOI: https://doi.org/10.47701/frasa.v6i2.5037

the ability to generate real-time evaluation data (Muhammad Yahya et al., 2023), enabling program administrators to efficiently monitor student progress and identify strengths and weaknesses in the curriculum (Martínez-Comesaña et al., 2023). Additionally, integrating AI in English proficiency assessment can help identify overall student performance trends, facilitating more data-driven decision-making regarding curriculum adjustments. With in-depth data analysis, AI allows for more precise adjustments to the teaching materials and learning methods used in the curriculum (Chen et al., 2022). The use of AI in curriculum evaluation also opens up opportunities for the development of more adaptive and responsive curricula. The data collected by AI can be utilized to design curriculum changes that are faster and more accurate, in line with the evolving needs of students and the demands of the English language profession.

CONCLUSION

The conclusion of this study is that the use of artificial intelligence (AI) in assessing English proficiency among students has proven to be effective and efficient, not only in individual assessment but also in supporting the evaluation of academic program curricula. AI is able to improve the objectivity and consistency of assessments, as well as significantly speed up the evaluation process. Additionally, AI provides a more accurate picture of curriculum achievement, enabling program administrators to track student progress more efficiently and identify areas needing improvement. The use of AI also opens up opportunities for the development of more adaptive and responsive curricula that address the needs of students and changes in professional demands. Thus, AI becomes a crucial tool in efforts to enhance the quality of education and achieve more effective learning outcomes.

However, this study has several limitations. First, it involved a limited number of respondents, which may not represent the entire population of students and lecturers using AI in English assessments. This can affect the generalizability of the findings. Second, the study focused more on the effectiveness and efficiency of AI in assessment, thus it has not deeply explored other aspects such as the long-term impact of AI usage on students' critical thinking or creativity skills. Third, external factors like technical conditions and the quality of data used by AI systems can also influence assessment results, but this has not been the primary focus of this study. Therefore, further research is needed to address these limitations and provide a more comprehensive understanding of AI's use in English proficiency assessment.

The findings of this study carry significant practical and theoretical implications for curriculum design and institutional policy. Theoretically, the study adds to the expanding conversation about AI in education by showing how it could make assessments more objective, efficient, and in line with program goals, while simultaneously pointing out the problems with judging complex language skills. In practice, the study gives instructors useful information. For example, it suggests that AI-based assessments might help create adaptive curricula that meet the needs of each student. This is especially important in teacher training programs where competency and teaching abilities are very important. The results show that institutions need to provide specialized training to make the most of AI, invest in infrastructure to make technological problems less of a problem, and set ethical norms to protect people's privacy.

FRASA: English Education and Literature Journal
Vol. 6 No. 2, September 2025

ISSN: 2807-8195
Available online at: https://ojs.udb.ac.id/
DOI: https://doi.org/10.47701/frasa.v6i2.5037

These consequences call for a balanced use of AI, making sure technology adds to rather than replaces human judgment in creating English education systems that are fit for the future.

However, the study also highlights the need for adequate technological readiness and training for both lecturers and students to optimize AI utilization. In addition, it is essential to continually monitor and evaluate the impact of AI on the teaching and learning process to ensure that this technology truly provides the expected benefits and does not create new inequalities in education.

REFERENCES

Albiladi, W. S., & Alshareef, K. K. (2019). Blended learning in english teaching and learning: A review of the current literature. *Journal of Language Teaching and Research*, *10*(2), 232–238. https://doi.org/10.17507/jltr.1002.03

An, X., Ching, ·, Chai, S., Li, Y., Zhou, Y., Shen, · Xi, Zheng, C., & Chen, · Mengyuan. (2023). Modeling English teachers' behavioral intention to use artificial intelligence in middle schools. *Education and Information Technologies*, *28*, 5187–5208.

Bahman, L., & Adrian, P. (2022). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford University Press.

Balakumar, P., Al-Qahtani, A., Shanmugam, K., Srividhya, P. ., & Sundram, K. (2023). In The Quintessence of Basic and Clinical Research and Scientific Publishing. In *How to Efficiently Write a Persuasive Discussion Section* (pp. 755–760).

Baxter, P., & Jack, S. (2008). Qualitative Case Study Methodology: Study Design and Implementation for Novice Researchers. *The Qualitative Report*, *13*(4), 544–559. https://doi.org/10.46743/2160-3715/2008.1573

Bin, Y., & Mandal, D. (2019). English teaching practice based on artificial intelligence technology. *Journal of Intelligent & Fuzzy Systems*, *3*(37), 3381–3391.

Chen, X., Zou, D., Xie, H., Cheng, G., & Liu, C. (2022). Two Decades of Artificial Intelligence in Education: Contributors, Collaborations, Research Topics, Challenges, and Future Directions. *Educational Technology and Society*, *25*(1), 28–47.

Chiu, T. K. F., Meng, H., Chai, C. S., King, I., Wong, S., & Yam, Y. (2022). Creation and Evaluation of a Pretertiary Artificial Intelligence (AI) Curriculum. *IEEE Transactions on Education*, *65*(1), 30–39. https://doi.org/10.1109/TE.2021.3085878

Delgado, H. O. K., de Azevedo Fay, A., Sebastiany, M. J., & Silva, A. D. C. (2020). Artificial intelligence adaptive learning tools: the teaching of English in focus. *Brazilian English Language Teaching Journal*, *11*(2), 1–19. https://doi.org/10.15448/2178-3640.2020.2.38749

El Shazly, R. (2021). Effects of artificial intelligence on English speaking anxiety and speaking performance: A case study. *Expert Systems*, *38*(3). https://doi.org/10.1111/exsy.12667

Evanini, K., Hauck, M. C., & Hakuta, K. (2017). *Approaches to automated scoring of speaking for K–12 English language proficiency assessments*. *1*, 1–11.

Fernandez, C. J. (2018). Behind a spoken performance: test takers' strategic reactions in a

FRASA: English Education and Literature Journal
Vol. 6 No. 2, September 2025

ISSN: 2807-8195
Available online at: https://ojs.udb.ac.id/
DOI: https://doi.org/10.47701/frasa.v6i2.5037

simulated part 3 of the IELTS speaking test. *Language Testing in Asia*, *8*(1). https://doi.org/10.1186/s40468-018-0073-4

Frost, K., Clothier, J., Huisman, A., & Wigglesworth, G. (2020). Responding to a TOEFL iBT integrated speaking task: Mapping task demands and test takers' use of stimulus content. *Language Testing*, *1*(37), 133–155.

Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, *9*(2), 113–132. https://doi.org/10.1080/15434303.2011.642041

Gayed, J. M., Carlon, M. K. J., Oriola, A. M., & Cross, J. S. (2022). Exploring an AI-based writing Assistant's impact on English language learners. *Computers and Education: Artificial Intelligence*, *3*(February), 100055. https://doi.org/10.1016/j.caeai.2022.100055

Huang, B. H., & Flores, B. B. (2018). The English Language Proficiency Assessment for the 21st Century (ELPA21). *Language Assessment Quarterly*, *15*(4), 433–442. https://doi.org/10.1080/15434303.2018.1549241

Kabudi, T., Pappas, I., & Olsen, D. H. (2021). AI-enabled adaptive learning systems: A systematic mapping of the literature. *Computers and Education: Artificial Intelligence*, *2*(March), 100017. https://doi.org/10.1016/j.caeai.2021.100017

Koraishi, O. (2023). Teaching English in the Age of AI: Embracing ChatGPT to Optimize EFL Materials and Assessment. *Language Education & Technology Journal*, *3*(1), 55–72.

Kyle, K., Crossley, S. A., & McNamara, D. S. (2016). Construct validity in TOEFL iBT speaking tasks: Insights from natural language processing. *Language Testing*, *3*(33), 319–340.

Manongga, D., Rahardja, U., Sembiring, I., Lutfiani, N., & Yadila, A. B. (2022). RETRACTED (Di Tarik) : Dampak Kecerdasan Buatan Bagi Pendidikan. *ADI Bisnis Digital Interdisiplin Jurnal*, *3*(2), 10-124-. https://doi.org/https://doi.org/10.34306/abdi.v3i2.792

Martínez-Comesaña, M., Rigueira-Díaz, X., Larrañaga-Janeiro, A., Martínez-Torres, J., Ocarranza-Prado, I., & Kreibel, D. (2023). Impact of artificial intelligence on assessment methods in primary and secondary education: Systematic literature review. *Revista de Psicodidáctica (English Ed.)*, *28*(2), 93–103. https://doi.org/10.1016/j.psicoe.2023.06.002

Mikalef, P., & Gupta, M. (2021). Artificial intelligence capability: Conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance. *Information and Management*, *58*(3), 103434. https://doi.org/10.1016/j.im.2021.103434

Muhammad Yahya, Wahyudi, & Akmal Hidayat. (2023). Implementasi Artificial Intelligence (AI) di Bidang Pendidikan Kejuruan Pada Era Revolusi Industri 4.0. *Seminar Nasional Dies Natalis 62*, *1*, 190–199. https://doi.org/10.59562/semnasdies.v1i1.794

Nazari, N., Shabbir, M. S., & Setiawan, R. (2021). Application of Artificial Intelligence powered digital writing assistant in higher education: randomized controlled trial. *Heliyon*, *7*(5), e07014. https://doi.org/10.1016/j.heliyon.2021.e07014

Owen, N., Shrestha, P. N., & Hultgren, A. K. (2021). Researching Academic Reading in Two Contrasting English as a Medium of Instruction Contexts at a University Level. *ETS Research Report Series*, *2021*(1), 1–28. https://doi.org/10.1002/ets2.12317

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),

FRASA: English Education and Literature Journal
Vol. 6 No. 2, September 2025

ISSN: 2807-8195
Available online at: https://ojs.udb.ac.id/
DOI: https://doi.org/10.47701/frasa.v6i2.5037

32–40. https://doi.org/10.25047/jlct.v1i2.3967

Schepman, A., & Rodway, P. (2023). The General Attitudes towards Artificial Intelligence Scale (GAAIS): Confirmatory Validation and Associations with Personality, Corporate Distrust, and General Trust. *International Journal of Human-Computer Interaction*, *39*(13), 2724–2741. https://doi.org/10.1080/10447318.2022.2085400

Swiecki, Z., Khosravi, H., Chen, G., Martinez-Maldonado, R., Lodge, J. M., Milligan, S., Selwyn, N., & Gašević, D. (2022). Assessment in the age of artificial intelligence. *Computers and Education: Artificial Intelligence*, *3*(August 2021). https://doi.org/10.1016/j.caeai.2022.100075

Tang, K. Y., Chang, C. Y., & Hwang, G. J. (2023). Trends in artificial intelligence-supported e-learning: a systematic review and co-citation network analysis (1998–2019). *Interactive Learning Environments*, *31*(4), 2134–2152. https://doi.org/10.1080/10494820.2021.1875001

Wang, B., Rau, P. L. P., & Yuan, T. (2023). Measuring user competence in using artificial intelligence: validity and reliability of artificial intelligence literacy scale. *Behaviour and Information Technology*, *42*(9), 1324–1337. https://doi.org/10.1080/0144929X.2022.2072768

Xu, Y., Liu, X., Cao, X., Huang, C., Liu, E., Qian, S., Liu, X., Wu, Y., Dong, F., Qiu, C. W., Qiu, J., Hua, K., Su, W., Wu, J., Xu, H., Han, Y., Fu, C., Yin, Z., Liu, M., … Zhang, J. (2021). Artificial intelligence: A powerful paradigm for scientific research. *Innovation*, *2*(4). https://doi.org/10.1016/j.xinn.2021.100179

Yang, Y., Wang, Q., Woo, L. H., & Quek, C. L. (2011). Using Facebook for teaching and learning: a review of the literature. *International Journal of Continuing Engineering Education and Life Long Learning*, *1*(21), 72–86.

Zahedi, K., & Shamsaee, S. (2012). Viability of construct validity of the speaking modules of international language examinations (IELTS vs. TOEFL iBT): Evidence from Iranian test-takers. *Educational Assessment, Evaluation and Accountability*, *24*(3), 263–277. https://doi.org/10.1007/s11092-011-9137-z

Zahedkazemi, E. (2015). Construct Validation of TOEFL-iBT (as a Conventional Test) and IELTS (as a Task-based Test) among Iranian EFL Test-takers' Performance on Speaking Modules. *Theory and Practice in Language Studies*, *5*(7), 1513. https://doi.org/10.17507/tpls.0507.27

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, *16*(1). https://doi.org/10.1186/s41239-019-0171-0