Role of Artificial Intelligence in Language Assessment

Omar Mohammed Ali Mohammed al Shaykhi

Assistant Professor -department of English language

-Faculty of Education Benghazi - University of Benghazi

دور الذكاء الاصطناعي في التقويم و القياس اللغوي

عمر محمد علي محمد الشيخي

أستاذ مساعد ـقسم اللغة الإنجليزية ـكلية التربية بنغازي ـجامعة بنغازي

omerkroby@gmail.com

Abstract

Language assessment and evaluation is crucial for employment, education, and language proficiency. Language evaluation has long employed human evaluators to rate and assess language competency in accordance with preset criteria. However, this manual assessment method has some limitations, including subjectivity, inter-rater variability, and scalability issues. The rapid advancement of AI technology has led to significant improvements in language assessment, producing more creative, accurate, and effective evaluation methods. The current study covers a wide range of subjects, including automated scoring and evaluation, benefits and advantages, challenges and considerations, and future directions. AI can assist language evaluation in achieving previously unheard-of levels of scalability and impartiality while accounting for ethical considerations. The goal of this article is to comprehend AI's impact on language assessment and guiding for further research and development in this dynamic field.

Keyword: artificial intelligence; language assessment; automated scoring

ملخص الدراسة

إن تقييم واخنبار إجادة اللغة أمر فائق الأهمية للتوظيف والتعليم. لقدُ استخدم تقييم اللغة منذ زمن طويل مُقيِّمين بشريين لتقييم وتقويم كفاءة وإجادة اللغة وفقًا لمعايير محددة مسبقًا. ومع ذلك، فإن طريقة التقييم اليدوية هذه بها بعض القيود، بما في ذلك الذاتية، والتباين بين المُقيِّمين، وقضايا قابلية التوسع. أدى التقدم السريع لتكنولوجيا الذكاء الاصطناعي إلى تحسينات كبيرة في تقييم اللغة، مما أدى إلى إنتاج طرق تقييم وقياس أكثر إبداعًا ودقة وفعالية. تغطي الدراسة الحالية مجموعة واسعة من الموضوعات، بما في ذلك التسجيل والتقييم الآلي، والفوائد والمزايا، والتحديات والاعتبارات، والاتجاهات المستقبلية. يمكن للذكاء الاصطناعي أن يساعد في تقييم اللغة في تحقيق مستويات غير مسبوقة من قابلية التوسع والنزاهة مع مراعاة الاعتبارات الأخلاقية. الهدف من هذه المقالة هو فهم تأثير الذكاء الاصطناعي على تقييم اللغة والتوجيه لمزيد من البحث والتطوير في هذا المجال الديناميكي.

الكلمات المفتاحية: الذكاء الاصطناعي، تقييم اللغة، التقييم الآلي.

Introduction

Language evaluation is one field that has advanced significantly as a result of the integration of Artificial Intelligence (AI) into many other fields. Although it is widely accepted that computer-based learning, data understanding, and managing complexity in difficult situations are all components of artificial intelligence (AI), it is crucial to remember that computers drive this process rather than being innate. According to **Cugurullo (2020)**, there isn't a single definition for artificial intelligence. AI has sparked a revolution with the rapid advancement of machine learning and natural language processing (NLP) techniques, as highlighted by **González-Calatayud et al. (2021),** and **Huang et al. (2022).**

In a variety of settings, including school, immigration, and the workplace, language evaluation is essential. Language tests have historically placed a great deal of reliance on human assessors, which frequently incorporated subjectivity and possible biases into the judgment process. Together with the growing availability of smaller, cheaper computers, these technologies had a significant impact on the capabilities that enabled Computer Adaptive Testing (CAT) **(Reckase, 2017).**

These systems may evaluate spoken or written language samples and deliver automated, reliable, and rapid scoring by utilizing NPL algorithms and machine learning models. AI-based language assessment provides a consistent and effective method of assessing language proficiency by removing the subjectivity of human judgment. **(Nardi & Ranieri, 2019)**. Moreover, they give test-takers instant feedback, facilitating focused language skill development and individualized learning opportunities. Additionally, a variety of linguistic characteristics, including grammar, vocabulary usage, fluency, coherence, and even sentiment analysis, can be analyzed by AI-driven systems. A more comprehensive assessment of language proficiency is made possible by this thorough evaluation, which picks up on subtleties that conventional tests could have missed. As a result, AI-based language assessment tools make it possible to depict a person's competency more accurately, encouraging impartial and equitable evaluation procedures **(Quansah, 2018)**.

Current Challenges in Language Assessment

Introduction.

By definition, language is an abstract phenomenon, and this abstraction adds to the difficulties in evaluating it. As the area of language evaluation develops, it must contend with both established issues and fresh ones brought about by the social, economic, and educational circumstances of the twenty-first century. A careful analysis of present assessment procedures is required due to the growing demands for high language proficiency among professionals and students as well as the requirement for accountability in language instruction. Due to its intrinsic abstraction, language is a very complex phenomenon, particularly when it comes to evaluation. The more abstract a notion is, the harder it is to quantify. This is seen in the wide-ranging subject of language assessment, where a number of recurring problems have been found. Even if a lot of problems have been solved by thorough research, as the area develops, new problems keep coming up. The field is lively and fascinating for those who are committed to it because of these problems, which cut across social, educational, and economic contexts. There are still gaps despite the fact that many research have addressed different aspects of language assessment. **(Iskandarova, 2024).**

Both classroom language assessment and high stakes accountability assessment are facing significant challenges due to the growing number of language learners in schools. How we will actually use or apply the knowledge we have acquired is the problem in classroom language evaluation. In order to achieve this, we must first create assessments that serve the goals of teaching and learning. Second, all of the teachers in language learning classrooms will need to receive training. How we will use the knowledge we have as language testers is the problem for accountability assessments.? To do this, we must create an assessment that evaluates students' performance not just in the language of instruction but also in a number of other subjects, like science and math. The assessment language must be distinct from the test-taker's home tongue. Another layer of difficulty is the growing use of language tests to assess applicants' eligibility for political asylum. **(Gan & Lam, 2022)**.

Dynamic Assessment

One of the more modern methods for evaluating language learners' development is dynamic assessment. One technique for performing a language assessment to ascertain a person's aptitudes and learning potential is called "dynamic assessment." It is widely regarded as a useful approach and is being utilized by educational psychologists worldwide more and more. It is frequently referred to as an active measurement method for the perception, understanding, reasoning, and problem-solving skills of the learner. Additionally, it has been shown to be very process-oriented and interactive. In fact, using it to evaluate students from diverse linguistic and cultural backgrounds is really advantageous. To clarify, it can be said that dynamic assessment has been shown to influence students by altering their cognitive processes and bridging the gap between cultural-linguistic evaluation and standard testing. (**Alsaadi, 2021**). **Poehner & Wang (2021)** have explained that the purpose of dynamic assessment is to help students succeed with their future assignments by negotiating with teachers, not only to help them finish a particular work.

A common definition of dynamic assessment (DA) is a strategy that concurrently incorporates teaching and assessment activities. As an example, DA depicts instruction and assessment as two sides of a single coin that are inseparable, meaning that effective instruction necessitates effective assessment, and effective assessment necessitates instruction. Thus, two complementary facets of methodology—instruction and assessment—should work together to produce genuine learning. Additionally, because the assessment approach stresses a person's capacity to acquire knowledge and skills that are conceptualized of the exam by a particular instruction, other research have characterized it as a test-teach retest model (**Tabatabaei & Bakhtiarvand, 2014**).

Mehri and Amerian (2015) have defined dynamic assessment as a method that combines evaluation and training to enhance learners' growth through suitable mediation. A organized mapping process using unknown objects to indicate pretense during a play is part of the Dynamic Assessment approach. This allows the learner to recognize the intended word and employ it in the discourse.

Alsaadi, (2021) have stated that teacher-student contact is necessary for dynamic evaluation. By changing the format, presenting more examples or trials, modeling a suitable approach for replies, or giving increasingly direct clues or prompts, the teacher tries to assist the student in creating paths when they are struggling to solve a problem or respond to a question. Additionally, by outlining the sequential steps that can be taken to establish dynamic assessment, they have given dynamic assessment methods in a very clear way as follow:

1. The examiner tests the learner working alone (static mode) to provide a measurement of skills on a task to establish a baseline.

2. the examiner provides a controlled protocol of assistance and instruction (dynamic mode) while the child works on a comparable task.

3. a post-test is given with an alternate form of the original measurement while the learner works alone (static mode) on the task.

4. The examiner compares the test and retest measurements to establish the learners' zone of proximal development.

5. The examiner analyzes the learner's performance on both product and process.

However, this form can be assumed to be complicated compared to the various forms of dynamic assessment during classes that teachers simply can practice to measure their students' comprehension and performance. Since dynamic assessment entails a type of active teaching and learning through perception, observation, and thought, it may be feasible to adjust the learner's cognitive processes and problem-solving styles to account for individual differences. There is no relationship between students' performance, learning potential, and social backgrounds, according to research done to identify the variables causing poor performance levels among students. These studies indicate that a lack of learning experiences, cultural differences, particular learning challenges, or traumatic life events are the main reasons why many students perform poorly on static examinations. Thus, it has been suggested that the fields of psychoanalysis and counselling employ dynamic assessment **(Rashidi and Bahadori Nejad, 2018).**

Forms and Types of Language Testing

1. Aptitude Tests

This refers to the ability of an individual to acquire knowledge. Tests of language aptitude evaluate a person's capacity to pick up new language abilities. Due to their nature, these tests are not language-specific and are more generic than the majority of other language examinations. Rather, they evaluate a person's capacity to pick up new language abilities efficiently and rapidly. To choose the best workers to enrol in language classes so they may help set up a new international branch or offer bilingual customer support, a company may administer an aptitude test **(Winna & Sabarun, 2023)**.

2. Diagnostic Tests

The purpose of diagnostic tests is to determine the current level of a person's proficiency in a certain domain, in this case, language. Diagnostic tests are usually administered at the beginning of a language learning program or course, as opposed to achievement and proficiency examinations. A diagnostic test typically consists of questions or tasks that are outside the limits of the test taker's knowledge and skills. The test's outcomes show a person's language proficiency's strong and weak points. Teachers can create lesson plans that address the gaps in a student's present abilities by using the findings of the student's diagnostic tests. Diagnostic tests are another tool that students can utilize to identify their areas of weakness on in order to reach a higher level of proficiency **(Lee, 2015)**.

3. Placement Tests

Diagnostic tests and placement tests are quite comparable. They are given prior to the start of a course or study program and are used for instructional purposes. The application is rather different in this instance. Placement tests are used by teachers and administrators to assign language learners to study groups or classes based on their skill levels. To decide whether a new French major has to complete introductory French courses or skip some and start with more advanced programs, a university may administer a placement test. Given that international students usually arrive with varying histories and degrees of English proficiency, placement tests are also a crucial component of university-level English language instruction. **(Fan & Jin, 2020).**

4. Achievement Tests

An achievement exam assesses a student's language proficiency to demonstrate the growth of their learning. Achievement exams, in contrast to diagnostic, aptitude, and placement tests, only cover material that the student should have already encountered during their academic career. At the end of the course or after a class has finished a particular chapter or unit, achievement assessments are usually administered. At the conclusion of the semester, a language instructor may administer a final test to assess how well a student has retained the material covered. Tests of achievement are usually graded and intended to show how well a language learner is doing in their studies **(Jin & Zhang, 2021).**

5. Proficiency Tests

A student's level of proficiency is assessed using components from the entire language, not simply the material that was discussed in class. It attempts to assess how well a student understands a language in comparison to other students who may have completed a different course or to particular linguistic domains identified by a language analysis. Assessing a learner's level of language progress is the goal of a proficiency exam. The International English Language Testing System (IELTS) and the Test of English as a Foreign Language (TOEFL) are two of the most well-known tests. Around the world, these tests are usually offered for a price. It is necessary to analyze the structure of linguistic competence in order to assess it. Any theory, hypothesis, or model that explains the world we live in is a structure. Candidates will have to finish tasks that show they can use English correctly in a specific scenario. Developing these missions and then doing research to confirm them involves a significant amount of time and money. **(Durairajan, 2016).**

An Overview of Artificial Intelligence Technologies

The goal of the multidisciplinary field of artificial intelligence (AI) is to create computer systems that are capable of carrying out tasks that typically call for human intelligence. It includes a number of subfields, such as deep learning, machine learning, and natural language processing (NLP). These technologies have been used extensively in language assessment, revolutionizing the way tests are administered and opening up new avenues for more accurate and efficient evaluation **(Zhai & Wibowo, 2023).**

Machine learning is a branch of AI that enables computer systems to learn from data and gradually enhance their functionality without explicit programming. Large databases of language samples can be used to train machine learning algorithms for language evaluation, which then automatically recognize patterns, linguistic traits, and performance indicators. It makes it possible to create AI models that assess language competency and produce accurate results **(Al-Abbas et al., 2023)**.

Deep learning is a branch of machine learning that models and simulates the structure and operation of the human brain using artificial neural networks. Large volumes of language data may be processed and analyzed by deep learning algorithms, enabling more intricate and subtle evaluations. Speech recognition, language production, and complicated language comprehension have all benefited greatly from this technology. (Yu & Deng, 2015)**.**

The study of how computers and human language interact is known as natural language processing, (NLP). It entails creating methods and algorithms to comprehend, evaluate, and produce voice or text in natural language. Because it makes it possible to process and analyze spoken responses, written essays, and other language samples,

natural language processing (NLP) has proven crucial to language evaluation. Sentiment analysis, syntactic parsing, and discourse analysis are examples of NLP approaches that can offer important insights regarding language ability and usage. **(Al-Abbas et al., 2023)**.

To evaluate and comprehend language samples, natural language processing techniques have also been incorporated into language evaluation. For instance, sentiment analysis can evaluate the sentiment or emotional tone conveyed in spoken or written comments. Analyses of textual complexity and readability can assess written texts' degree of difficulty and guarantee that learners are using the right language. The field has changed as a result of the use of AI in language assessment, which provides automated scoring and assessment of intelligent tutoring systems **(Al-Abbas et al., 2023)**.

In many different fields, artificial intelligence (AI) has become a disruptive force that is changing industries and how jobs are carried out. AI technologies are transforming conventional pedagogical approaches in the field of education, especially in English Language Teaching (ELT), and providing creative ways to improve language learning experiences. This introduction lays the groundwork for examining the significant effects of AI in ELT, emphasizing how it may be used to tailor education, create immersive learning environments, and meet the various needs of students. Additionally, AI technologies provide real-time evaluation and feedback, giving students prompt direction on vocabulary, grammar, pronunciation, and other language abilities. AI improves the efficacy and efficiency of language training by offering individualized guidance and reinforcement, whether through virtual language assistants, interactive chatbots, or automated grading systems. **(Sharma et al., 2024)**.

In the field of educational technology research, artificial intelligence (AI) is beginning to contribute to the development of improved methods for educational assessment. AI can assist in the analysis of enormous volumes of educational data by using algorithms or neural networks to complete tasks that call for intelligent behavior. There are three key advantages to using AI to assist in creating more reliable educational tests. First, AI can assist meet the increasing need for high-quality evaluations by powering automated processes that create fair, legitimate, and trustworthy assessments quickly. Second, intelligent machines are better equipped to handle the task of creating and delivering customized tests that can promote learning as educational evaluations become more automated. Third, unlike human intelligence, artificial intelligence (AI) can be scaled up and utilized to create and administer extensive tests in a variety of global contexts **(Hamon et al., 2022)**.

An Integration of AI in Language Assessment

In the 'Artificial Intelligence in Educational evaluation' section, the use of AI technology in educational evaluation is examined. As a spine that determines whether and how well learning has occurred, educational assessments are essential to education. The evaluation methods used today, however, are antiquated, uninteresting, inefficient for higher-order thinking skills, and unable to keep up with the rapid changes in student instruction and the classroom. As a result, the field of educational evaluation requires modernization, and one of the technologies propelling this change is artificial intelligence (AI). The amount of research interest in this topic is growing quickly **(Chen et al., 2020)**.

Integrating AI technologies in language assessment has brought significant advancements and benefits to the field. Assessment procedures can be automated and streamlined by AI-based language assessment systems, yielding more unbiased, trustworthy, and scalable results. AI models may evaluate language samples, recognize linguistic traits, and offer automated scoring and feedback by utilizing machine learning, natural language processing, and deep learning approaches **(Al-Abbas et al., 2023)**.

AI-powered automated scoring systems can assess spoken comments, written essays, and other language samples. These systems are capable of analyzing linguistic elements such as argumentation, coherence, grammar, and vocabulary and producing unbiased, reliable assessments. Automated scoring helps learners grow and improve their language skills by saving time and providing instant feedback. **(Sevcikova, 2018)**.

Intelligent tutoring systems, which provide language learners with personalized instruction and feedback, are another application of AI in language assessment. These systems can identify students' areas of strength and weakness, modify the curriculum to suit their needs, and more. Intelligent tutoring systems that make use of AI technology can provide engaging and dynamic learning environments that enhance students' language skills **(Shi et al, 2018)**.

1. Automated Scoring and Evaluation

A key component of language assessment that makes use of Artificial Intelligence (AI) technologies to examine and evaluate a range of language skills is automated scoring and evaluation. It includes assessment of grammar and vocabulary, evaluation of spoken language, and automated essay scoring, among other areas. These AI-powered language assessment programs have transformed the evaluation process by providing reliable, impartial, and effective outcomes **(Mizumoto & Eguchi, 2023)**.

2. Automated Essay Scoring

AI algorithms are used in automated essay scoring to assess and grade written essays. The use of AI in language evaluation has drawn a many interests because of its capacity to process vast numbers of essays and deliver prompt, reliable comments. Several NLP techniques are used by automated essay scoring systems to examine the language elements of essays, such as vocabulary, grammar, coherence, and reasoning. To forecast grades for new, unseen writings, the systems employ machine learning algorithms that have been trained on enormous datasets of manually evaluated essays. This method improves productivity, lessens human bias, and gives students prompt feedback on their writing **(Ramesh & Sanampudi, 2021)**.

It is crucial that AI-based grading systems are dependable and efficient. By contrasting the outcomes produced by AI systems with those from human assessors, researchers are able to verify the correctness of automated essay scoring. Research evaluates the dependability of automated essay scoring by comparing human and AI-generated results (Kumar & Boulanger, 2020). In order to guarantee uniformity and equity in the evaluation of essays, research also looks into how AI-based scoring aligns with accepted rubrics and assessment frameworks **(Kaldaras & Haudek, 2022)**.

3. Spoken Language Proficiency Assessment

Another area in which AI has significantly advanced is the assessment of spoken language abilities. When evaluating spoken language ability, artificial intelligence (AI) tools like speech recognition and natural language processing are essential. By converting spoken language into text, speech recognition algorithms allow for the investigation of a variety of linguistic features, including intonation, fluency, and pronunciation. The transcribed text is then processed using natural language processing tools to assess the coherence, vocabulary usage, and correctness of spoken responses **(Zhai & Wibowo, 2023)**.

There are many advantages to employing AI for spoken language evaluation, such as consistency, objectivity, and scalability. AI effectively assesses a huge number of spoken language samples by automating the evaluation process. By continuously using predetermined rating standards, AI systems remove human biases and guarantee evaluation impartiality. Furthermore, learners can discover and address areas for growth in their spoken language skills thanks to the instant feedback that AI-based spoken language evaluation gives them **(Xu et al, 2021)**.

4. Grammar and Vocabulary Assessment

In language assessments, AI technologies have also been used to evaluate vocabulary and grammar skills. AI algorithms are used by automated error detection and repair systems to examine spoken or written language samples and find vocabulary and grammar mistakes. To identify mistakes and provide fixes, these systems use statistical methodologies, rule-based approaches, and machine-learning techniques. AI-based grammar and vocabulary evaluation has many benefits. It makes it possible to evaluate language proficiency effectively and impartially, especially when it comes to spotting typical mistakes and giving students focused comments. By providing recommendations for error correction and alternate word choices, automated error detection and correction systems can assist learners in increasing the accuracy of their language and broadening their vocabulary. These artificial intelligence (AI)-powered resources enhance conventional grammar and vocabulary training by offering students individualized assistance throughout their language learning process **(Yu & Deng, 2015)**.

Automated essay scoring, spoken language assessment, grammar and vocabulary assessment, and other AI-based language assessment applications are all included in automated scoring and evaluation. These programs' efficiency, objectivity, and consistency have revolutionized the appraisal process. AI-based solutions manage high exam volumes, give prompt feedback, and help students advance their language proficiency. The dependability and efficiency of these AI-driven technologies are being improved by ongoing research and development in automated scoring and evaluation, which empowers teachers and students in language assessment **(Al-Abbas et al., 2023)**.

Benefits and Challenges

The field of educational evaluation is undergoing a significant transformation due to the development and availability of Artificial Intelligence (AI) models. There is a great chance to develop creative solutions for efficient, forward-thinking, and contextual educational assessments by integrating AI technology, generative tools, and digital platforms. AI tools can promote inclusive and equitable educational outcomes by streamlining and offering affordable alternative digital techniques. Additionally, intelligent systems can produce new methods of engagement for both teachers and students, offering new contexts and intricacies to mold questions and responses in

situations of global citizenship. AI may provide a variety of symbol-interrelation questions and provide fair educational results in the home or in the classroom. The generated questions, pass rate, and safeguards against attempts at correct answers are all greatly impacted by context and tools. Beyond privacy and academic integrity policymaking concerns, AI technology tools can foster diversity in the educational field through social influence, exposure, and interactions **(George & Wooden, 2023).**

However, approaches to engagement, assessment design, and evaluating digital methodology may shift as a result of the enthusiasm and concern surrounding AI tools. Educational institutions are becoming more conscious of and concerned about issues pertaining to academic integrity, access equity across socioeconomic backgrounds, and reevaluating assessment evaluation techniques. At various levels of educational institutions, a variety of initiatives, awareness-raising campaigns, and actions are being seen. When implementing AI technologies in educational institutions, personal domains, and age groups, there are two impending challenges: developing skilled individuals and maintaining academic integrity. Whether AI-generated questions should form the basis of assessment is a significant philosophical and evaluative concern for education (**Rane, 2024**).

By using AI to continuously enhance educational evaluations, educational measurements that are more in line with the evolving requirements of a widening and changing spectrum of educational activities and processes can be created. Modern societies are becoming more and more dependent on new technologies, and outside of the more traditional and permanent formal and structured settings of education (such as schools and universities), educational processes are dominated by either machines or human-centered technology-based activities. The need for educational assessments is increasing in this context due to the variety of teaching, learning, and instructional design opportunities offered by the quickly expanding fields of distance education, Massive Open Online Courses (MOOCs), learning analytics, Intelligent Tutoring Systems (ITS), and others. Better evaluations within evolving educational processes still require more study and work to ensure fairness, validity, and dependability **(Kuleto et al., 2021).**

Future Directions

Due to the introduction of AI, educational assessment is currently undergoing a significant revolution. Diverse trends in the advancement and use of AI in educational evaluation are reflected in the body of current work. These patterns may provide the foundation for more research and studies on the use of AI in evaluating students' knowledge and abilities. Recent work has identified the lack of understanding of these developments as an urgent concern. Delineating the trends and areas of concentration of AI applications, particularly in the field of educational assessment and evaluation, is becoming more and more necessary in order to enable the unique application of AI in education. The success of applications may be better understood through trend analysis, which will help identify directions for the development of AI in education and the provision of AI-based assessment tools for educational systems. This does not imply that teachers will be replaced by AI, though, as formative evaluations and the human element will always be important. However, a thorough grasp of the potential of AI tools and instruments will be essential to the efficacy, creativity, validity, and fairness of formal evaluations inside educational systems **(Huang et al., 2023).**

An improvement in the quality of immediate feedback mechanisms made possible by AI technology is one possible future development that has been proposed.

Learning requires feedback, but it can be challenging to deliver timely, high-quality input. AI tools make it possible to gather information about students as they complete assignments so that feedback can be given immediately, adaptively, and in accordance with current procedures, either immediately before or immediately after the assignment, improving learning chances. **(Chen et al., 2020)**.

The field of language assessment has benefited from improvements in speed, accuracy, availability, objectivity, and reliability due to advancements in AI technology. Though more work and creativity are required, ongoing technology advancements will enable our structures to be expanded to capture increasingly complex facets of language proficiency and automatically score (**Van Moere & Downey, 2016**).

AI is a useful tool for teaching foreign languages. Artificial intelligence is playing an increasingly significant role in foreign language learning; digital services created with AI technology are incorporated into instructional strategies, and in the future, AI may support teachers in their pedagogical endeavors by taking over some of their duties. All things considered, artificial intelligence holds great promise for improving language proficiency. AI has the potential to greatly improve our language skills and language learning procedures in the future if innovation, ethical considerations, and ongoing technological growth are properly balanced (**Qiao & Zhao, 2023).**

Conclusion

Artificial intelligence is poised to revolutionise educational assessment by offering scalable and efficient solutions. Artificial intelligence is an extremely useful instrument for improving language proficiency. Through its advanced and sophisticated algorithms, analytical capabilities, and personalised approaches, AI transforms how we assess, understand, and improve language. Its contribution lies in the precision of language skill assessment, modifying tests to individual needs, efficiency in evaluation, and providing personalised guidance for improvement. The integration of artificial intelligence into educational assessment marks a transformative shift in how we evaluate language proficiency. By offering scalable and efficient solutions, AI not only enhances the precision of skill assessment but also tailors evaluations to meet individual needs. Advanced algorithms facilitate a deeper understanding of language abilities, allowing for targeted improvements that were previously unattainable. This personalized guidance fosters an environment where learners can thrive, ensuring that evaluation is both efficient and effective. As we embrace these innovative technologies, the future of language assessment promises to be more adaptive and insightful, ultimately reshaping the educational landscape for generations to come.

REFERENCES

Al-Abbas, H. U. A. I., Halim, H. H., & Nurjati, N. N. (2023). Harnessing the use of artificial intelligence in language assessment: A systematic comprehensive review. Tell-Us Journal, 9(3), 723-745.

Alsaadi, H. M. A. (2021). Dynamic Assessment in Language Learning: An Overview and the Impact of Using Social Media. English Language Teaching, 14(8), 73-82.

Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. Ieee Access, 8, 75264-75278.

Cugurullo, F. (2020). Urban artificial intelligence: From automation to autonomy in the smart city. Frontiers in Sustainable Cities, 2, 38.

Durairajan, G. (2016). Assessing and testing learners' language proficiency. Language and Language Teaching, 5(1), 52-56.

Fan, J., & Jin, Y. (2020). Standards for language assessment: Demystifying university-level English placement testing in China. Asia Pacific Journal of Education, 40(3), 386-400.

Gan, L., & Lam, R. (2022). A review on language assessment literacy: Trends, foci and contributions. Language Assessment Quarterly, 19(5), 503-525.

George, B., & Wooden, O. (2023). Managing the strategic transformation of higher education through artificial intelligence. Administrative Sciences, 13(9), 196.

González-Calatayud, V., Prendes-Espinosa, P., & Roig-Vila, R. (2021). Artificial intelligence for student assessment: A systematic review. Applied sciences, 11(12), 5467.

Hamon, R., Junklewitz, H., Sanchez, I., Malgieri, G., & De Hert, P. (2022). Bridging the gap between AI and explainability in the GDPR: towards trustworthiness-by-design in automated decision-making. IEEE Computational Intelligence Magazine, 17(1), 72-85.

Huang, F., & Sun, L. (2023). Examining the Roles of Technology in Sustaining Language Teaching and Learning. Sustainability, 15(24), 16664.

Huang, X., Zou, D., Cheng, G., Chen, X., & Xie, H. (2023). Trends, research issues and applications of artificial intelligence in language education. Educational Technology & Society, 26(1), 112-131.

Iskandarova, G. (2024). Current issues in language assessment and language assessment research and its implication. Baltic Journal of Legal and Social Sciences, (3), 228-231.

Jin, Y., & Zhang, L. J. (2021). The dimensions of foreign language classroom enjoyment and their effect on foreign language achievement. International Journal of Bilingual Education and Bilingualism.

Kaldaras, L., & Haudek, K. C. (2022). Validation of automated scoring for learning progression-aligned next generation science standards performance assessments. In Frontiers in Education (Vol. 7, p. 968289). Frontiers Media SA.

Kuleto, V., Ilić, M., Dumangiu, M., Ranković, M., Martins, O. M., Păun, D., et al. (2021). Exploring opportunities and challenges of artificial intelligence and machine learning in higher education institutions. Sustainability, 13(18), 10424.

Lee, Y. W. (2015). Diagnosing diagnostic language assessment. Language Testing, 32(3), 299-316.

Mehri, E., & Amerian, M. (2015). Challenges to dynamic assessment in second language learning. Theory and Practice in Language Studies, 5(7), 1458.

Mizumoto, A., & Eguchi, M. (2023). Exploring the potential of using an AI language model for automated essay scoring. Research Methods in Applied Linguistics, 2(2), 100050.

Nardi, A., & Ranieri, M. (2019). Comparing paper-based and electronic multiple-choice examinations with personal devices: Impact on students' performance, self-efficacy and satisfaction. British Journal of Educational Technology, 50(3), 1495-1506.

Poehner, M. E., & Wang, Z. (2021). Dynamic assessment and second language development. Language Teaching, 54(4), 472-490.

Qiao, H., & Zhao, A. (2023). Artificial intelligence-based language learning: illuminating the impact on speaking skills and self-regulation in Chinese EFL context. Frontiers in Psychology, 14, 1255594.

Quansah, F. (2018). Traditional or Performance Assessment: What is the Right Way in Assessing Leaners? Research on Humanities and Social Sciences; Vol 8, No 1 (2018); 21-24.

Ramesh, D., & Sanampudi, S. K. (2022). An automated essay scoring systems: a systematic literature review. Artificial Intelligence Review, 55(3), 2495-2527.

Rane, N. (2024). Enhancing the quality of teaching and learning through Gemini, ChatGPT, and similar generative Artificial Intelligence: Challenges, future prospects, and ethical considerations in education. TESOL and Technology Studies, 5(1), 1-6.

Rashidi, N., & Bahadori Nejad, Z. (2018). An investigation into the effect of dynamic assessment on the EFL learners' process writing development. Sage Open, 8(2), 2158244018784643.

Reckase, M. D. (2017). A tale of two models: Sources of confusion in achievement testing. ETS Research Report Series, 2017(1), 1-15.

Sevcikova, B. L. (2018). Human versus Automated Essay Scoring: A Critical Review. Arab World English Journal, 9(2).

Sharma, S., Krishnamaraju, J. M. R., & Divakaran, P. (2024). Impact of Artificial Intelligence in English Language Teaching. Migration Letters, 21, 1606-1612.

Shi, G., Lippert, A. M., Shubeck, K., Fang, Y., Chen, S., Pavlik, P., ... & Graesser, A. C. (2018). Exploring an intelligent tutoring system as a conversation-based assessment tool for reading comprehension. Behaviormetrika, 45, 615-633.

Tabatabaei, S., & Bakhtiarvand, M. (2014). Application of Dynamic Assessment in Second and Foreign Language Teaching. International Journal for Teachers of English, 4(3), 1-14.

Van Moere, A., & Downey, R. (2016). 21. Technology and artificial intelligence in language assessment. Handbook of second language *assessment*, 341-358.

Winna, W., & Sabarun, S. (2023). The Language Assessment In Teaching-Learning English. DIAJAR: Jurnal Pendidikan dan Pembelajaran, 2(4), 413-419.

Xu, J., Jones, E., Laxton, V., & Galaczi, E. (2021). Assessing L2 English speaking using automated scoring technology: examining automarker reliability. Assessment in Education: Principles, Policy & Practice, 28(4), 411-436.

Yu, D. & Deng, L. (2015). Automatic Speech Recognition A Deep Learning Approach. London: Springer-Verlag London.

Zhai, C., & Wibowo, S. (2023). A systematic review on artificial intelligence dialogue systems for enhancing English as foreign language students' interactional competence in the university. Computers and Education: Artificial Intelligence, 4, 100134.