E-ISSN: 2985-5616
Volume 02, Nomor 03, Desember 2023

J-SES
Journal of Science, Education and Studies

ARTIFICIAL INTELLEGENCE (AI) IN ASSOCIATION WITH LANGUAGE ASSESSMENT

***Ratnayanti[1], Rosarina Poppy Handayani[2], Sri Wahyuni[3], Nunung Nurjati[4]**

Universitas PGRI Adi Buana Surabaya, Indonesia, ratnayanti164@gmail.com: https://orchid.org/0000-0001-5529-5182
Universitas PGRI Adi Buana Surabaya, Indonesia,rosarinapoppy33@gmail.com: https://orchid.org/0000-0002-7025-5354
Universitas PGRI Adi Buana Surabaya, Indonesia,wahyuni.sri298@gmail.com : https://orchid.org/0009-0006-6021-3296
Universitas PGRI Adi Buana Surabaya, Indonesia, nunung.nurjati@unipasby.ac.id: https://orchid.org/0000-0003-4877-7663

Corresponding Author:
***Ratnayanti[1]**
Universitas PGRI Adi Buana Surabaya, Indonesia
ratnayanti164@gmail.com

Article History

Received: 23-10-2023
Revision: 23-10-2023
Acceptance: 20-11-2023
Published: 04-12-2023

Abstrak: *Kecerdasan buatan (AI) telah menjadi topik perbincangan di berbagai bidang, termasuk pendidikan. Dengan kemajuan teknologi. Makalah ini mengeksplorasi penggunaan kecerdasan buatan (AI) dalam penilaian bahasa. Studi ini menganalisis penelitian terkini mengenai penilaian siswa berbasis AI dan potensi dampaknya terhadap pendidikan yang menerapkan AI, serta metode penilaian kualitas yang digunakan dalam studi penilaian bahasa berbasis AI. Makalah ini menyoroti potensi AI dalam membantu guru dalam penilaian siswa dan memberikan umpan balik. Hal ini juga berkontribusi pada pemahaman tentang bagaimana AI dapat dimanfaatkan dalam penilaian bahasa di bidang pendidikan dan memanfaatkan potensi penuh AI dalam pendidikan. Makalah ini memberikan wawasan berharga mengenai kondisi penelitian terkini dalam penilaian pendidikan berbasis AI. Temuan ini menunjukkan bahwa pendekatan berbasis AI berpotensi mengatasi keterbatasan metode penilaian tradisional, seperti keterbatasan waktu dan akurasi.*

Katakunci: *Penilaian Bhs. Inggri; AI; Analisa; Integrasi*

Abstract: *Artificial Intelligence (AI) has been a topic of discussion in various fields, including education. with the advancement of technology. This paper explores the use of artificial intelligence (AI) in language assessment. The study analyzes the current state of research in AI-based student assessment and its potential impact on education where AI is*

E-ISSN: 2985-5616
Volume 02, Nomor 03, Desember 2023

applied, and the quality assessment methods used in AI-based language assessment studies. The paper highlights the potential of AI in assisting teachers in student assessment and providing feedback. It also contributes to the understanding of how AI can be utilized in language assessment in education and harnesses the full potential of AI in education. The paper provides valuable insights into the current state of research in AI-based educational assessment. The findings suggest that AI-driven approaches have the potential to overcome limitations in traditional assessment methods, such as time contraints and accuracy.

Keywords: *English Language Assessment, AI, Analysis, Explore, Integrate*

Introduction

The incorporation of Artificial Intelligence (AI) into various sectors has led to transformative advancements, and one area that has particularly progressed is language evaluation. While there is a consensus that artificial intelligence (AI) involves the integration of computer-based learning, comprehension of data, and handling complexity in challenging scenarios, it's important to note that this process is not inherent but rather driven by computers. According to Curugullo (2020), AI lacks a singular definition. With the swift evolution of natural language processing (NLP) and machine learning methodologies, AI has brought about a revolution, as highlighted by Chen et al. (2020), González-Calatayud et al. (2021), and Huang et al. (2022). This study takes an in-depth look at the novel applications of AI in language assessment, investigating its potential to improve the precision, efficiency, and impartiality of language proficiency evaluations.

Language assessment plays a crucial role in various contexts, ranging from education to immigation and employment. Traditionally, language assessments heavily relied on human evaluators, which often introduced subjectivity and potential biases into the evaluation process. The capabilities that made Computer Adaptive Testing (CAT) possible were driven by the influence of these technologies in combination with the rising availability of smaller, more affordable computers, stated Reckase, (2017). Computer-based language assessment can be more accurate and efficient than traditional methods of testing if it is designed properly, Chalhoub-Deville and Deville (1999); Chapelle, (2001). However, with the emergence of AI technologies, researcher and practitioners have started harnessing the power of algorithms and computational models to develop more objective and reliable language assessment tools,

One prominent application of AI in language assessment is automated scoring systems. The belief that a computer has been programmed to "do what humans do" is one widespread (mis)perception of automated scoring, Ericsson and Haswell (2006). Instructors' evaluations and third-party evaluations of writing samples not included in tests. Machines do not behave like people when automated scoring is used. Instead, It takes advantage of the fact that computers can be programmed to identify and rate writing and speaking abilities. It then combines and weights those elements in a multidimensional space to determine which specific features and their weightings are most accurate at predicting a person's score.

Leveraging NPL algorithms and machine learning models, these systems can analyze written or spoken language samples and provide automated, consistent, and timely scoring. By eliminating the subjectivity of human judgement, Alsubai *et al* (2021), AI-based language assessment offers a standardized and efficient approach to evaluate language proficiency, as Jin (2022); Nardi &

Ranieri, (2019) said. Moreover, they provide immediate feedback to test takers, enabling targeted language skill improvement and personalized learning experiences. Futhermore, AI-driven systems can analyze a range of linguistic features, such as grammar, vocabulary usage, fluency, coherence, and even sentiment analysis. This comprehensive evaluation allows for a more holistic assessment of language skills, capturing nuanced aspects that might have been overlooked in traditional assessments Quansah, (2018). Consequently, AI-based language assessment tools enable a more accurate representation of an individual proficiency, promoting fair and unbiased evaluation processess.

Another noteworthy application of AI in language assessment in the development of intelligence tutoring systems. These systems utilize AI algorithms to provide adaptive and personalized language instruction to learners, Watson *et al* (2021). By analizing individual learner's strengths, weaknesses, and learning patterns, intelligent tutoring systems can offer tailored exercises, materials, and feedback, fostering an optimized learning experience. This personalized approach not only enhances language acquisition but also supports learners in achieving their specific language goals.

Despite the immense potential of AI in language assessment, challenges and ethical considerations remain. The ethical use of AI algorithms, data privacy, and transparancy in decision-making are critical aspects that need to be addressed. Furthermore, it is crucial to continually validate and refine AI-based assessment tools to ensure their reliability, validity, and alignment with established language proficiency frameworks, David (2021).

In this paper, we bring the examination of AI and language assessment, incorporating imaginative and pragmatic viewpoints provided by specialists in the domain. Our intention is to delve into AI;S revolutionary capabilities in language assessment, thereby enriching the ongoing discourse regarding the prospective evolution of language evaluation and its implications for education, migration, and employment. Together, let us commence on this voyage to reveal the captivating progressions and possible hurdies that emerge at thr crossroads of artificial intelligence and language assessment.

Educational Setting For Language Assessment

Language assessment has traditionally relied on human evaluators to measure individuals' language proficiency levels. However, the subjective nature of human assessment and the limitations associated with manual grading have prompted researchers and practitioners to explore the integation of AI technologies in language assessment, currently stated by Iryna (2023). This setting aims to investigate the diverse applications of AI in language assessment, including scoring systems, intelligent tutoring

systems, and adaptive testing platforms.

Automated scoring systems

First of AI-Based on language assessment is Diagnostic accuracy instruments Jayakuma (2022). These instruments are employed to evaluate the precision of AI-driven diagnostic pattern like the quality of algorithms and input data, the practicality of real-world clinical use and the overal applicability of the algorithms. And it is the prominent applications of AI in language assessment is the development of automated scoring systems (Chapelle and Cho, 2010). These systems utilize natural language processing (NPL) techniques, machine learning algorithms, and linguistic features to evaluate and score language performance. Researchers have extensively examined the reliability and validity of automated scoring systems across various language skills, such as writing, speaking, and listening. The finding demostrate that these systems can provide reliable and consistent evaluations, often comparable to human raters. Futhermore, automated scoring systems offer the advantages of instant feedback, scalability, and cost- effectiveness, thus enabling efficient language assessment on a large scale.

Burr (2022) on his argument of Proficiency Levels. Machine learning and natural language processing methods are applied to generate proficiency scales according to a specified benchmark. Linguistic models are subsequently utilized to directly gauge the complexity of items for computer-adaptive testing, negating the requirement for costly preliminary trials involving human participants.

Automated Essay Grading. AI algorithms are employed for the automatic evaluation and scoring of essays. This process involves using techniques from natural language processing to analyze the content, structure, and coherence of the essays. As mentioned by Ramesh (2020).

Translation Evaluation Standards Chao (2020). AI-driven models such as ChatGPT possess the capacity to assess translations and measure their quality in real-time scenarios. These models utilize transformer-based architectures to analyze and compare translations. Framework for Linguistic Quality Assessment. Advanced frameworks for appraising language quality are employed to ensure clear and comprehensive quality assessment in AI-powered language evaluation systems.

Intelligent Tutoring Systems

AI-based intelligent tutoring systems have gained significant attention in language assessment research. These systems have gained significant attention in language assessment research. These systems leverage machine learning algorithms and personalized feedback mechanisms to provide tailored language instruction and assessment. Intelligent tutoring systems can analyze learners' performance, identify areas of improvement, and

offer adaptive learning materials accordingly. Several studies have demonstrated the effectiveness of intelligent tutoring systems in enhancing learners' language skills, motivation, and engagement. Moreover, these systems can generate comprehensive learner profiles, enabling educators to track progress and provide targeted interventions as said by P. Sedlmeie (2001).

Adaptive Testing Platforms.

AI-powered adaptive testing platforms have revolutionized the field of language assessment by dynamically adjusting the difficulty level test items based on learners' responses. These platforms employ Item Response Theory (IRT) models and algorithms to optimize assessment process. Adaptive testing provides a more precise estimation of learners' abilities and reduces the time required for assessment. It also ensures that test takers are presented with item that are neither too easy nor too difficult, resulting in more accurate and informative assessments Lord, (1970); Wainer (2000); Van der Linden and Glas (2010

Benefits and Challenges

The integration of AI in language assessment offers several benefits as stated by Van Moere, Alistair Downey, Ryan (2016), including enhanced objectivity, increased efficiency, immediate feedback provision, and personalized learning experiences. However, there are challenges associated with AI implementation, such as ensuring fairness, addressing the lack of transparency in AI algorithms, and considering the ethical implications of automated decision-making in high-stakes language assessments. These bchallenges require careful consideration to ensure the ethical and equitable use of AI in language assessment practices.

This literature review highlights the significant advancements made in the application of AI in language assessment. Automated scoring systems, intelligent tutoring systems, and adaptive testing platforms have demonstrated their potential in improving the accuracy, efficiency, and personalization of language assessment processes. As AI technologies continue to evolve, further research is needed to address the challenges and explore additional opportunities for integrating AI in language assessment. By leveraging the strengths of AI and human expertise, language assessment can become more reliable, effective, and learner-centered in the digital era.

Potential Future direction of AI in Language Assessment

As outlined by Zhenglin Zhang et al. (2020), the domain of AI within language assessment is swiftly developing, with numerous potential paths it could potentially follow in the future. The manner in which language proficiency is gauged has already been transformed due to the prevalent access to computers and the internet (Alderson 2000). Yet, it's apparent that this demand will continue to escalate. Here are a handful of potential avenues to explore.

Adaptive and personalized assessment said Peng (2019) that Ai could be used to create adaptive language assessments that dynamically adjust the difficulty of questions based on the test taker's responses. By analyzing patterns in the test taker's answer and using machine learning algorithms, AI could tailor the assessment to their skill level, providing a more accurate and personalized evaluation of their language abilities.

And there is Roukos (2008) who argues that Natural language understanding. AI models are continually improving their ability to understand and generate human language. Future developments in natural language understanding could enable more sophisticated language assessments. AI systems could analyze a test taker's responses not only for correctness but also for deeper understanding, coherence, and pragmatics, providing more comprehensive evaluations of language proficiency.

Ross (2020) concerns on Multimodal Assessment. Language assessment traditionally focuses on written or spoken language, but AI could facilitate the development of multimodal assessments that incorporate other forms of communication, such as gestures, facial expressions, and even virtual reality environments. By integrating multiple modalities, AI could capture a more holistic view of language abilities and assess a wider range of skills, such as communication in real-world contexts.

Daouda (2023) asserts on Continuous and Real-Tme Assessment. AI technology could enable continuous and real-time language assessment in various contexts. For example : AI-powered language assistants or chatbots could assess a person's language proficiency during conversational interactions, providing immediate feedback and targeted recommendations for improvement. This approach could be particularly useful for language learning and training purposes.

Different to Sun (2023, he focuses on Ethical Considerations. In order to assure the ethical use of these cutting-edge technologies, UNESCO (2021) is working on the creation of a worldwide framework to direct uses and applications of AI. To ensure human rights and dignity, we must weigh the many benefits against the hazards, evil uses, and divisions that we must foresee. As AI language Assessment becomes more prevalent, there will be a need to address ethical considerations. Fairness, bias, privacy, and security are important issues that need to be carefully addressed to ensure that AI assessments are reliable, transparent, and unbiased. Future directions in AI language assessment should prioritize these ethical considerations and work towards creating trustworty and equitable assessment systems.

It's important to note that these potential directions are speculative and based on current trends and advancements in AI. The actual future of AI in language assessment will depend on technological

advancements research breakthoughs, and the needs and priorities of the field.

Emphasizing The Need for Ethical Considerations & Human AI Collaboration

Emphasizing the need for ethical considerations and human AI collaboration is crucial in today's rapidly advancing technological landscape, Boni (2021). As artificial intelligence (AI) continues to evolve and integrate into various aspects of our lives, it is essential to ensure that its development and deployment are guided by ethical principles and human values. Here are some key reasons why ethical considerations and human-AI collaboration are necessary.

Accountability and Transparency. Raji et al., (2020) emphasizes that ethical considerations help hold AI systems and their developers accountable for their actions. By establishing clear guidelines and standards, we can ensure that AI technologies and transparent and understanable, allowing for scrutiny and accountability when needed. This transparency in crucial in building trust between human and AI systems.

Fairness and Bias Mitigatian Ashwathy (2021) argues that AI systems have the potential to perpetuate and amplify existing biases present in the data they are trained on. Ethical considerations demand that AI algorithms be developed with fairness in mind, mitigating biases and ensuring equitable treatment across different groups of individuals. Human involvement is necessary to detect and correct biases in AI systems, as a human judjementand emphaty are vital in addressing complex societal issues.

Privacy and Data Protection Quach (2022). As AI relies heavily on data, ethical considerations necessitate safeguarding individuals' privacy and protecting their personal information. Collaboration between humans and AI can help establish mechanism for responsible data collection, usage, and storage, ensuring that privacy rights are respected and protected.

Joseph (2014) talks about Human-Centered Design. AI systems should be designed with humans in mind. By involving humans in the design and development process, we can ensure that AI technologies align with human needs, values, and preferences. Human-AI collaboration can result in more user-friendly and accessible systems that enhance human capabilities rather than replace them.

Unintended Consequences and safety. Ethical consideration prompt us to anticipate and address the potential unintended consequences of AI-deployment, Suckling (2021). Collaboration between humans and AI can help identify and mitigate risks associated with AI technologies, ensuring safety and preventing harm to individuals or society as a whole.

As Ghonim (2022) stated that decision-Making and Value Alignment. AI systems are increasingly being involved in decision-making processes

that have significant societal implications. Ethical considerations call foe aligning AI decision-making with human values, ensuring that decisions made by AI systems are understantable, justifiable, and in line with human moral frameworks. Human involvement is vital in guiding and validating these decisions.

According to Ertemel (2021) Socioeconomic Impacts. Ethical considerations encompass the broader socioeconomic impacts of AI deployment. Human-AI collaboration can help address concerns such as job displacement and ensure that the benefits of AI technologies are distributed wquitably, minimizing potential negative effects on marginalized communities.

Exploring for Educational Setting

The incorporation of Artificial Intelligence (AI) into language evaluation has brought about a revolutionary change in the domain. This transformation presents novel possibilities for enhancing the precision, efficiency, and impartility of gauging language proficiency, as argued by Jia (2021). The primary objective of this research is to investigate the utilization of AI in language assessment. This analysis seeks to provide a more profound comprehension of the merits, limitations, and potential enhancement of AI-centered language evaluation mechanisms. The exploration aims to assemble these discoveries to establish a fundamental basis for diverse research endeavors, as the offer insights into the progression of knowledge, sources of inspiration for policy formulation, act as a catalyst for fresh concepts, and serve as a valuable research compass within specific areas of study. Evaluation of Translation excellence : AI-driven model like ChatGPT have the capability to appraise translations and gauge their caliber in live situations. These models employ transformer-based structures to examine and juxtapose translation.

Following an extensive investigation, employing the methodologies of thorough exploration, data compilation, and analysis, it is essential that additional research and refinement of AI systems for language assessment take place to overcome limitations and fully unlock their capabilities. The data collection phase of the study underscores the potential of AI applicatios in language assessment, presenting opportunities to enhance the learning process, provide individualized feedback, and elevate the impartiality of evaluations.

The application of artificial intelligent (AI) in language assessment has gained significant attention in recent years. Language assessment refers to the evaluation of an individual's proficiency in a particular language, typically focusing on areas such as speaking, writing, listening, and reading skills. AI technologies have the potential to revolutionize langguage assessment by providing automated and adaptive solutions that are efficient, objective, and scalable. The followings are some key point to consider in the result and discussion of AI application in language assessment.

E-ISSN: 2985-5616
Volume 02, Nomor 03, Desember 2023

Automated Scoring AI can be used to develop automated scoring systems for assessing language proficiency. Improvements in automated scoring of a wider variety of attributes may lead to more precise assessments of student abilities. Software, for instance, is available to assess how tones are used in languages like Mandarin (Cheng 2012). Tepperman and Nava (2011), Chen and Zechner (2011), and Tepperman, Kazemzadeh, and Narayanan (2007) research on the application of automated assessments of spoken prosody, rhythm, and intonation in speech shows promise to boost the reliability of nativeness estimates. Similar developments that will shed light on written language proficiency are also in the works.

It is not surprising that current research has concentrated on the use of automated scoring technology to rate productive language abilities like speaking and writing. Improved technology makes testing materials more accessible, eases the practical restrictions on administering and scoring tests, imposes standardization of delivery, enforces objectivity and dependability of scoring, and promotes scalability of use. In the literature, claims about the value of automated scoring are supported (e.g., Balogh et al. 2012; Downey et al. 2010). These systems utilize machine learning algorithms to analyze linguistic features and patterns in written or spoken responses, assigning scores based on predefined criteria. By automating the scoring process, AI can provide faster and consistent evaluation, reducing the need for human raters and improving efficiency. Automated Scoring in language assessment refers to the use of artificial intelligence (AI) and Natural Language Processing (NPL) technologies to automatically evaluate and score language proficiency or performance in various tests or assessments.

Adaptive Testing AI can enable adaptive testing, where the difficulty level of questions or tasks dynamically adjusts based on the test taker's performance. Adaptive systems use AI algorithms to select items that are most appropriate for the test taker's proficiency level, allowing for more accurate and personalized assessment. This approach enhances the precision of measuring language skills and provides a tailored experience for each individual.

Natural Language Processing (NPL) NPL is a branch of AI that focuses on the interaction between computers and human language. NPL technique can be applied in language assessment to analyze and understand text or speech, enabling automated comprehension and feedback. AI-powered system can detect grammatical error, evaluate vocabulary usage, assess coherence and cohesion in writing and even provide suggestion for improvement.

Language Generation AI can also be utilized to generate language based content for assessment purposes. For instance : AI algorithms can create writing prompts, reading passages, or spoken dialogue simulations that reflect real world language use. This allows for generation of diverse and

engaging test materials that accurately measure language skills across various contexts.

Ethical Considerations The use of AI in language assessment raises ethical concerns that need to be carefully addressed. Fairness and bias issues should be considered to ensure that AI systems do not discriminate against certain group based on language background, dialect, or cultural nuances. Transparency and explainability of AI algorithms are also importan to gain trust and acceptance from test takers, educators, and other stakeholders.

Supplementing Human Expertise While AI can automate ceratain aspects of language assessment, human expertise remains essential. AI system should be designed to complement human judgement rasther than replace it enterely. Combining the strengths of AI technology with human assessment can result in more accurate and comprehensive language evaluations.

It's important to note that the field of AI in language assessment is still evolving, and there are ongoing research and development effort to refine and improve the application mentioned above. Collaboration between AI experts, language assessment specialists, and educators is crucial to ensure the effective integration of AI technology into language assessment practices.

Summary

The author comes to the summary that According to the author's definition of artificial intelligence (AI), it is "an artifact able to acquire information on the surrounding environment and make sense of it, in order to act rationally and autonomously even in uncertain situations" González-Calatayud (2021). In conclusion, the application of artificial intelligence (AI) in language assessment has shown great promise and potential. Through its advanced capabilities, AI has provided innovative solutions to longstanding challenges in language assessment, such as objectivity, scalability, and efficiency. By leveraging natural language processing, machine learning, and deep learning techniques, AI systems can effectively evaluate and analyze various linguistic aspects, including grammar, vocabulary, fluency, and coherence.

The integration of AI in language assessment has brought numerous benefits to both test takers and educators. AI-based assessment tools offer personalized feedback and adaptive learning experiences, allowing individuals to identify their strengths and weakness and tailor their learning strategies accordingly. Moreover, AI systems can assess language proficiency in a timely manner, reducing the waiting time for test result and enabling faster feedback loops for learners. This accelerates the language learning process and facilitates targeted intervention for improvement.

The application of AI in language assessment has transformed the field, providing innovative solutions that address the traditional assessment challenges. However, while AI has

demostrated its potential in language assessment, it is crucial to acknowledge certain limitations ande ethical considerations. The development and implementation of AI systems should prioritize transparency, fairness, and bias mitigation. Careful attention must be given to the quality and diversity of training data to avoid perpetuating biases and ensuring equitable assessment practices. Moreover, human involvement and expertise remain essential in interpreting AI-generated results and making informed decisions about language proficiency.

In this journal, we have explored the vast lanscape of artifical intelligence (AI) application in language assessment. Looking ahead, the future of AI in language assessment holds exciting prospects. Continued advancements in AI technology, including natural language understanding, context awareness, and multimodal analysis, will futher enhance the accuracy and reliability of language assessment tools. The integration of AI with other merging technologies, such as virtual reality and augmented reality, may offer immersive language assessment experiences, enabling authentic and interactive evaluations.

References

Alderson, J. C. (2000). Assessing reading. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511732935

Alamri, H.A., Watson, S. & Watson, W. Learning Technology Models that Support Personalization within Blended Learning Environments in Higher Education. *TechTrends* 65, 62–78 (2021). https://doi.org/10.1007/s11528-020-00530-3

Ashwathy Ashokan, Christian Haas, Fairness metrics and bias mitigation strategies for rating predictions, Information Processing & Management, Volume 58, Issue 5, 2021, 102646, ISSN 0306-4573, https://doi.org/10.1016/j.ipm.2021.102646. (https://www.sciencedirect.com/science/article/pii/S0306457321001369)

Boni, M. (2021). The ethical dimension of human–artificial intelligence collaboration. *European View*, *20*(2), 182–190. https://doi.org/10.1177/17816858211059249

Burr Settles, Geoffrey T. aFlair, Masato Hagiwara; Machine Learning–Driven Language Assessment. *Transactions of the Association for Computational Linguistics* 2020; 8 247–263. doi: https://doi.org/10.1162/tacl_a_00310

Chao Han (2020) Translation quality assessment: a critical methodological review, The Translator, 26:3, 257-273, DOI: 10.1080/13556509.2020.1834751

Chen, X., Xie, H., Zou, D., & Hwang, G.-J. (2020). Application and theory gaps during the rise of Artificial Intelligence in Education. Computers

and Education: Artificial Intelligence, 1, 100002. https://doi.org/10.1016/j.caeai.2020.100002

Cugurullo, F. Urban Artificial Intelligence: From Automation to Autonomy in the Smart City. Front. Sustain. Cities 2020, 2, doi:10.3389/frsc.2020.00038.
Chalhoub-Deville, Micheline & Craig Deville. 1999. Computer adaptive testing in second language contexts. Annual Review of Applied Linguistics 19. 273-299.

Chapelle, Carol. 2001. Computer applications in second language acquisition: Foundations for Teaching, testing, and research. Cambridge, UK: Cambridge University Press.

Chapelle, Carol. 2003. English language learning and technology: Lectures on applied linguistics in the age of information and communication technology. Amsterdam: John Benjamins.

Chapelle, Carol & Dan Douglas. 2006. Assessing language through computer technology. Cambridge, UK: Cambridge University Press.

Chapelle CA and Chung Y-R (2010). The Promise of NLP and speech processing technologies in language assessment. Language Testing, 27(3): 301–315.

David J. Hand, Shakeel Khan, Validating and Verifying AI Systems, Patterns, Volume 1, Issue 3, 2020, 100037, ISSN 2666-3899, https://doi.org/10.1016/j.patter.2020.100037.

Daouda Kamissoko, Blaszho Nastov & Matthieu Allon, (2023). Improved Model for Continuous, Real-Time Assessment and Monitoring of The Resilience of System Based on Multiple Data Sources and takeholders, Structure and Infrastructure Engineering, 19:8, 1122-1137, DOI : 10.1080/15732479.2021.2009883

Ericsson, P. F., & Haswell, R. H. (Eds.). (2006). Front Matter. In *Machine Scoring of Student Essays: Truth and Consequences* (p. [i]-[iv]). University Press of Colorado. https://doi.org/10.2307/j.ctt4cgq0p.1

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

Glas, C. A. W., and Verhelst, N. D. (1995). "Testing the rasch model," in *Rasch Models Foundations, Recent Developments, and Applications*, eds G. H. Fischer and I. W. Molenaar (New York, NY: Springer Verlag), 69–95.

Ertemel, Adnan Veysel Karadayı, Talha Makaritou, Paraskevi. (2021) P 75- 89. INVESTIGATING THE SOCIO-ECONOMIC CONSEQUENCES OF

ARTIFICIAL INTELLIGENCE: A QUALITATIVE RESEARCH. Vol. 7

Ghonim, M.A., Khashaba, N.M., Al-Najaar, H.M, and Khashan, M.A (2022), 'Strategic Alignment and Its Impact on Decision Effectiveness: A comprehensive Model'. International Journal of

Emerging Markets. Vol. 17 No. 1. Pp.198-218. https://doi.org/10.1108/IJOEM-04-2020-0364

Iryna Alves & Sofia M. Lourenço (2023) Subjective performance evaluation and managerial work outcomes, Accounting and Business Research, 53:2, 127-157, DOI: 10.1080/00014788.2021.1959292

Jayakumar, S., Sounderajah, V., Normahani, P. *et al.* Quality assessment standards in artificial intelligence diagnostic accuracy systematic reviews: a meta-research study. *npj Digit. Med.* **5**, 11 (2022). https://doi.org/10.1038/s41746-021-00544-y

Jia-Cing Liang, Gwo-Jen Hwang, Mei-Rong Alice Chen & Darmawansah Darmawansah (2021) Roles and research foci of artificial intelligence in language education: an integrated bibliographic analysis and systematic review approach, Interactive Learning Environments, DOI: 10.1080/10494820.2021.1958348

Joseph Giacomin (2014) What Is Human Centred Design?, The Design Journal, 17:4, 606-623, DOI: 10.2752/175630614X14056185480186

Huang, J., Saleh, S., & Liu, Y. (2021). A Review on Artificial Intelligence in Education. Academic Journal of Interdisciplinary Studies, 10(3), 206. https://doi.org/10.36941/ajis-2021-0077

Jin, L. (2022). Design of English Writing Assessment System Based on AI Technology in food and agriculture sectors. Journal of Commercial Biotechnology, 25(4). https://doi.org/10.5912/jcb1252

Lord, F. M. (1970). "Some test theory for tailored testing," in *Computer-Assisted Instruction, Testing, and Guidance*, ed W.H. Holtzma (New York, NY: Harper and Row), 139–183.

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. https://doi.org/10.1111/bjet.12644

Peng, H., Ma, S. & Spector, J.M. Personalized adaptive learning: an emerging pedagogical approach enabled by a smart learning environment. *Smart Learn. Environ.* **6**, 9 (2019). https://doi.org/10.1186/s40561-019-0089-y

P. Sedlmeier, Intelligent Tutoring Systems, Editor(s): Neil J. Smelser,

Paul B. Baltes, International Encyclopedia of the Social & Behavioral Sciences, Pergamon, 2001, Pages 7674-7678, ISBN 9780080430768, https://doi.org/10.1016/B0-08-043076-7/01618-1. (https://www.sciencedirect.com/science/article/pii/B0080430767016181)

Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020). In proceedings of the 2020 conference on fairness, accountability, and transparency. In *Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing* (pp. 33–44). ACM Digital Library.

Ramesh, D., Sanampudi, S.K. An automated essay scoring systems: a systematic literature review. *Artif Intell Rev* **55**, 2495–2527 (2022). https://doi.org/10.1007/s10462-021-10068-2 Reckase, M.D. (2017), A Tale of Two Models: Sources of Confusion in Achievement Testing. ETS Research Report Series, 2017: 1-15. https://doi.org/10.1002/ets2.12171

Ross, J., Curwood, J. S., & Bell, A. (2020). A multimodal assessment framework for higher education. *E-Learning and Digital Media*, *17*(4), 290–306. https://doi.org/10.1177/2042753020927201

Roukos, S. (2008). Natural Language Understanding. In: Benesty, J., Sondhi, M.M., Huang, Y.A. (eds) Springer Handbook of Speech Processing. Springer Handbooks. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-49127-9_31

Van Moere, Alistair Downey, Ryan. 2016. P 341 - 358- 978-1-61451-624-8 Technology and Artificial

Intelligence in Language Assessment. DOI: 10.1515/9781614513827-023

Suckling James, Hoolohan Claire, Soutar Iain, Druckman Angela. Unintended Consequences: Unknowable and Unavoidable, or Knowable and Unforgivable? Frontiers in Climate. Vol.3. 2021 https://www.frontiersin.org/articles/10.3389/fclim.2021.737929. DOI:10.3389/fclim.2021.737929 ISSN=2624-9553

Sun, F., Ye, R. Moral Considerations of Artificial Intelligence. Sci & Educ 32, 1-17 (2023). https://doi.org/10.1007/s11191-021-00282-3

Syrowatka, A., Kuznetsova, M., Alsubai, A. *et al.* Leveraging artificial intelligence for pandemic preparedness and response: a scoping review to identify key use cases. *npj Digit. Med.* 4, 96 (2021) https://doi.org/10.1038/s41746-021-00459-8

Van Moere, A. and Downey, R. (2016). Technology and artificial intelligence in language assessment In Dina Tsagari & Jayanti Banerjee (Eds.), Handbook of Second Language Assessment, De Gruyter Mouton.

https://doi.org/10.1515/9781614513827-023

Van der Linden, W. J., and Glas, C. A. W. (eds.). (2010). *Elements of Adaptive Testing.* NewYork, NY: Springer. doi: 10.1007/978-0-387-85461-8

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. https://www.iiste.org/Journals/index.php/RHSS/article/view/40787

Quach, S., Thaichon, P., Martin, K.D. *et al.* Digital technologies: tensions in privacy and data. *J. of the Acad. Mark. Sci.* **50**, 1299–1323 (2022). https://doi.org/10.1007/s11747-022-00845-y

Zhenglin Zhang et al, Artificial Intelligence-Powered Acoustic Analysis System for Dysarthria Severity Assessment, *Advanced Intelligent Systems* (2023). DOI: 10.1002/aisy.202300097