APPLICATION OF GENERATIVE ARTIFICIAL INTELLIGENCE TO CLASSROOM ASSESSMENT PRACTICES

[1]IMASUEN, KENNEDY & [2]ADEOSUN, P. K.

[1*2]*Institute of Education, Department of Educational Evaluation and Counselling Psychology, University of Benin, Benin City.*

kennedy.imasuen@uniben.edu

Abstract

Generative Artificial Intelligence (GenAI) refers to Artificial Intelligence (AI) systems that create new content, such as text, images, and audio, by leveraging large datasets and advanced machine learning models. Particularly in the educational sector, Generative artificial intelligence has revolutionized classroom assessment practices by enabling the creation of personalized, adaptive, and diverse assessment formats. These artificial intelligence-driven tools generate assessment items such as multiple-choice questions, essay prompts, and problem-solving tasks tailored to individual student needs. This paper explores the impact of generative artificial intelligence on classroom assessments, highlighting its ability to provide adaptive assessments, reduce cheating risks, and offer real-time feedback. By examining the opportunities generative artificial intelligence presents, the paper discusses practical applications, strategies for implementation, and methods to involve students in artificial intelligence-enhanced assessment design. The discussion culminates with suggestions for future directions in leveraging generative artificial intelligence to optimize classroom assessments and improve learning outcomes.

Keywords: Artificial intelligence, generative artificial intelligence, assessment, formative assessment

Introduction

Generative Artificial Intelligence refers to artificial intelligence systems capable of creating new content, including text, images, audio, and video, by learning from large datasets. These systems often rely on transformer models such as Generative Pre-trained Transformers (GPT) and Bidirectional Encoder Representations from Transformers (BERT)**.** GPT is a deep learning model designed to generate human-like text by predicting the next word in a sequence. It is pre-trained on extensive datasets and fine-tuned for applications like chatbots, content creation, and language translation. On the other hand, BERT is specifically designed for natural language processing (NLP) tasks. Unlike traditional models, it processes words in relation to their surrounding context (both before and after), enabling a deeper understanding of the meaning in a text. Both GPT and BERT are based on the

transformer architecture, which allows them to achieve state-of-the-art performance in NLP tasks.

In education, generative artificial intelligence has the potential to revolutionize assessment by creating multiple-choice questions, essay prompts, and problem-solving tasks. By analyzing student performance data, artificial intelligence can generate questions that align with specific learning objectives and difficulty levels. It can also personalize assessments by tailoring questions to individual student proficiency, adapting test difficulty in real-time, and enhancing test security by providing unique question sets. Beyond assessments, generative artificial intelligence supports personalized learning by adjusting instructional content based on student needs. It also enhances formative assessments, enabling educators to track student progress more effectively. However, while AI-driven educational tools offer significant benefits, concerns regarding ethics, bias, and data privacy must be carefully managed to ensure responsible implementation.

According to Vaswani et al. (2017), generative artificial intelligence is a subset of artificial intelligence, involving models capable of generating text, images, and other data forms based on learned patterns. The integration of Generative artificial intelligence in educational settings, particularly in classroom assessment practices, presents both opportunities and challenges. This paper aims to explore the potential impact of GenAI on classroom assessment practices, examining how these technologies can enhance or hinder educational outcomes, and it will be presented under the following subheadings:

The Concept of Generative Artificial Intelligence (AI); Generative Artificial Intelligence in Education; Classroom Assessment; Impact of generative Artificial Intelligence on Classroom Assessment Practices; Practical Applications of Generative Artificial Intelligence in Classroom Assessment; Types of Assessment in Relation to Generative Artificial Intelligence Usage; Examples of Generative AI Applications in Classroom Assessment, Strategies for Implementing Generative Artificial Intelligence in Assessments; Strategies to Involve Students in Designing Artificial Intelligence-Enhanced Assessments; Current Challenges and Future Directions; Conclusions and Recommendations

The Concept of Generative Artificial Intelligence (AI)

Artificial Intelligence (AI) is defined as "the science and engineering of making intelligent machines, brilliant computer programs. It is related to the similar task of using computers to understand human intelligence, but artificial intelligence does not have to confine itself to biologically observable methods" (Russell & Norvig, 2021, p. 1). This broad definition encompasses a wide range of AI applications, including self-driving vehicles, automatic language translation, facial recognition, virtual assistants, recommendation systems, and generative artificial intelligence. Generative artificial intelligence refers to artificial intelligence applications that can generate new content, such as text, images, or code. Examples

of generative AI include chatbots like ChatGPT and Bard, which use natural language processing (NLP) to provide human-like responses to questions. While the field of artificial intelligence encompasses far more than just generative artificial intelligence, the rapid emergence of chatbots in education makes this a particularly relevant area of focus. It is important to acknowledge that the field of artificial intelligence is rapidly evolving, and new platforms and resources will continue to be developed (Russell & Norvig, 2021). Generative artificial intelligence refers to artificial intelligence systems that can generate new data that resembles the data they were trained on. These systems use sophisticated machine learning models, particularly deep learning models, to understand and replicate patterns in the training data. (IBM, 2024).

Generative AI operates through several fundamental phases, including training, tuning, and generation (MIT News, 2023). It is widely applied across various domains, such as text generation, where it creates articles, stories, and even code, with tools like OpenAI's ChatGPT serving as notable examples. In the field of image and video creation, generative AI produces realistic visuals, making it valuable in entertainment and advertising. Additionally, it contributes to audio and music production by composing music and generating lifelike speech. In software development, it plays a crucial role in assisting with writing and debugging code, enhancing efficiency and accuracy.

Generative Artificial Intelligence in Education

Generative artificial intelligence technologies have seen rapid adoption in educational contexts, driven by their potential to transform teaching and learning processes. According to a report by Cornell University (2023), generative artificial intelligence tools can support various educational activities, from content creation to personalized learning experiences. These tools offer innovative ways to engage students and provide tailored feedback, thereby enhancing the overall learning experience (Cornell University, 2023). Perkins et al. (2024) highlight the need for a balanced approach that considers both the benefits and potential drawbacks of generative artificial intelligence. The Artificial Intelligence Assessment Scale (AIAS) proposed by Perkins and colleagues provides a framework for integrating generative artificial intelligence into educational assessments ethically and effectively. This framework emphasizes transparency, fairness, and the importance of aligning generative artificial intelligence usage with educational objectives (Perkins et al., 2024).

Classroom Assessment

Assessment plays a crucial role in education, extending beyond the common belief that classroom assessment merely provides evidence of learners mastering the content taught. As societal expectations and the goals and objectives of schools evolve, new approaches to classroom assessment are emerging. The understanding

of learning and learning outcomes is also shifting, moving away from the notion that assessment is solely the teacher's responsibility and is typically conducted at the end of a programme, course, or unit. Nowadays, assessment is viewed as an ongoing activity that helps teachers adjust their teaching methods in real-time. Students are now active participants in the assessment process, serving as self-assessors and playing a significant role in their evaluation. Various levels of assessment exist, including classroom-level assessments, external examinations, national assessments, and international assessments (Afemikhe, 2023).

Classroom assessment has traditionally been categorized as formative and summative. However, according to Imasuen and Adeosun (2022), the purposes of classroom assessment include assessment as learning (AAL), assessment of learning (AOL), and assessment for learning (AFL). These purposes are linked to formative and summative assessments. Formative assessment is further divided into assessment for learning and assessment as learning, while summative assessment is seen as an assessment of learning. These different purposes necessitate distinct planning and raise various quality issues. Assessment For Learning plays a pivotal role in student learning by gathering and analyzing data that students and teachers use to identify the student's current status, their goals, and the most effective path to achieve those goals (Budiyono & Mardiyana, 2019; Scheelenkans et al., 2021). It also serves as an investigative tool for teachers to learn about their students and refine their teaching methods (Afemikhe, 2023). It helps teachers understand what students know and can do, their feelings about the learning process, and any biases or gaps they may have (Sherlyare Hendri et al., 2019; Tjendami et al., 2019). Teachers use assessment for learning to boost students' motivation and commitment to learning (Hidayat et al., 2023). As Cohen et al. (2020) state, when teachers focus on assessment, they can transform the classroom culture into one of student success.

Practical Applications of Generative AI in Classroom Assessment
The integration of generative artificial intelligence in classroom assessments can lead to more dynamic and adaptive evaluation methods. Generative artificial intelligence can evaluate student essays by analyzing content, grammar, and coherence, providing consistent and unbiased grading. Tools like Grammarly and Turnitin's Revision Assistant offer real-time feedback on writing quality and originality (Dwivedi et al., 2023). Artificial Intelligence can generate customized learning materials and assessments based on individual student performance and learning styles, addressing unique needs and ensuring appropriate levels of challenge and support. Platforms like Dream Box and Knewton use artificial intelligence to adapt lessons in real time (Bala et al., 2023). Generative artificial intelligence can create interactive simulations and virtual labs for hands-on learning experiences, assessing student performance in real time and providing immediate feedback (Beckingham et al., 2024).

Artificial Intelligence-driven tools can conduct formative assessments by generating

quizzes and practice tests tailored to the curriculum, providing instant feedback to help students learn from their mistakes. Platforms like Quizlet and Kahoot use Artificial Intelligence to create engaging and adaptive quizzes (Perkins et al., 2024). Generative artificial intelligence can assist in language learning by generating conversational practice scenarios and assessing pronunciation, grammar, and vocabulary usage. Tools like Duolingo and Rosetta Stone incorporate artificial intelligence to provide personalized language learning experiences and assessments (Dwivedi et al., 2023). Artificial Intelligence can facilitate peer review processes by generating guidelines and rubrics for students to evaluate each other's work, promoting collaborative learning and critical thinking. Tools like Peergrade and Feedback Fruits use artificial intelligence to streamline peer assessment and ensure constructive feedback (Bala et al., 2023). Generative artificial intelligence can create adaptive tests that adjust the difficulty level based on the student's responses, ensuring appropriately challenging assessments that accurately measure student proficiency. The GRE and GMAT examinations use adaptive testing algorithms to tailor the difficulty of questions in real-time (Beckingham et al., 2024). Artificial Intelligence can analyze student performance data in real-time to identify trends and patterns, helping educators make informed decisions about instructional strategies and interventions. Platforms like Edmodo and Canvas use artificial intelligence to provide insights into student engagement and achievement (Perkins et al., 2024). Beckingham et al. (2023) observed that Generative artificial intelligence can be used to create more engaging and interactive assessment tasks. However, they also caution against over-reliance on these technologies, which can lead to issues such as academic misconduct and reduced critical thinking skills among students (Beckingham et al., 2023).

Types of Classroom Assessment in Relation to Generative Artificial Intelligence Usage

Generative Artificial Intelligence (GenAI) can be integrated into various types of classroom assessments to enhance educational practices.

Formative assessments: Generative artificial intelligence can be used during the learning process to monitor student progress and provide ongoing feedback. It can generate personalized quizzes, practice tests, and interactive exercises that adapt to the student's learning pace and needs. For example, platforms like Quizlet and Kahoot use Artificial Intelligence to create adaptive quizzes that provide immediate feedback, helping students identify areas for improvement (Perkins et al., 2024). Generative artificial intelligence can also be used for Summative assessments. It can be used to evaluate student learning at the end of an instructional period, such as final examinations or end-of-term projects. It can also assist in creating and grading these assessments. Automated essay scoring systems, like those used by Turnitin, can evaluate written responses for content, grammar, and coherence, ensuring consistent and unbiased grading (Dwivedi et al., 2023).

In terms of diagnostic assessments generative artificial intelligence is used to identify students' existing knowledge, skills, and learning gaps before instruction begins. Generative artificial intelligence can analyze student responses to diagnostic tests and generate detailed reports that highlight strengths and areas needing improvement. This helps educators tailor their instruction to meet individual student needs (Bala et al., 2023). In Peer assessment generative artificial intelligence helps students to evaluate each other's work. It can facilitate this process by generating rubrics and guidelines for peer reviews, ensuring that feedback is constructive and aligned with learning objectives. Tools like Peergrade use artificial intelligence to streamline peer assessment and provide insights into the quality of peer feedback (Beckingham et al., 2024).

Generative artificial intelligence is also helpful in adaptive assessments as it can adjust the difficulty of questions based on the student's responses, by providing a personalized assessment experience. It can create adaptive tests that dynamically change in real time to match the student's proficiency level. This approach is used in standardized tests like the GRE and GMAT, which employ adaptive testing algorithms to tailor question difficulty (Beckingham et al., 2024). In terms of performance-based assessments, generative artificial intelligence can be used to assess how students demonstrate their knowledge and skills through practical tasks, such as projects, presentations, or experiments. Generative artificial intelligence can create virtual simulations and interactive scenarios that allow students to engage in hands-on learning experiences (Dwivedi et al., 2023). Generative artificial intelligence can also provide tools for self-assessment by generating reflective prompts and personalized feedback based on student input. This helps students develop metacognitive skills and take ownership of their learning (Perkins et al., 2024).

Generative artificial intelligence enhances students' critical awareness of the technology and improves their writing skills (Hadley & Ardito, 2024), is used by students as an assistant for revising and homework, while educators use it for content production and personalizing learning materials (Clos & Chen, 2024). Both groups used generative artificial intelligence as an accessibility aid to rephrase sentences and explain concepts. Key concerns included equity in access, clarity of usage rules, and job displacement. Generative artificial intelligence can create interactive and personalized learning environments in medical education (BMC Medical Education, 2024).

Current Challenges in Artificial Intelligence and Classroom Assessment

One of the major challenges in integrating artificial intelligence into classroom assessments is ensuring data privacy and security. AI systems require large amounts of student data, raising concerns about potential misuse, breaches, and the need for stringent protection measures. Balancing effective AI utilization with strong data safeguards remains an ongoing issue (MIT News, 2023). Another challenge is bias in

AI algorithms, as these systems can reflect and perpetuate biases present in their training data. This could lead to unfair assessments influenced by factors such as race, gender, or socioeconomic background (Binns, 2022).

Teacher training and readiness also present significant obstacles. Many educators lack the necessary skills and knowledge to effectively incorporate AI tools into their assessment practices. Providing professional development opportunities and access to resources is crucial to addressing this gap (VanLehn, 2023). Additionally, AI-driven assessments may lack human judgment, as they are unable to fully account for emotional, social, and contextual factors that teachers naturally consider when evaluating students. This limitation can result in assessments that do not fully capture a student's learning experience (Miller, 2022).

Future Directions for Artificial Intelligence in Classroom Assessment

AI has the potential to transform assessment by enabling personalized learning experiences. It can tailor evaluations and provide real-time feedback that adapts to each student's progress, enhancing individualized education (MIT News, 2023). A promising approach is the development of hybrid human-AI systems, where AI offers data-driven insights while educators apply their expertise to ensure a more comprehensive and well-rounded assessment (VanLehn, 2023).

Ensuring ethical AI development is another key focus for the future. AI-driven assessment systems must be transparent, fair, and explainable, with ongoing efforts to minimize biases and uphold ethical standards (Binns, 2022). Gamified and adaptive assessments represent another exciting possibility, allowing AI to create interactive, engaging learning environments where assessments adjust dynamically based on a student's learning needs and motivation. By implementing these advancements, AI can support more effective, inclusive, and meaningful classroom assessments (Miller, 2022).

Strategies for Implementing Generative Artificial Intelligence in Classroom Assessments

Implementing generative artificial intelligence effectively requires a thorough understanding of the technology, clear policies and guidelines, AI-enhanced assessments, ethical considerations, and a collaborative learning environment. Educators should begin by familiarizing themselves with AI tools, their functions, benefits, and limitations. Participating in workshops, webinars, and training sessions can help them stay updated on developments and best practices in AI for education.

Clear policies and guidelines are essential to ensure responsible AI use in assessments. Defining acceptable practices and making students aware of these guidelines helps maintain academic integrity. Measures such as requiring students to

document their use of AI tools and submit drafts showing their thought processes can promote transparency and accountability. AI can also enhance assessments by enabling the creation of formative evaluations that provide immediate feedback, helping students improve continuously. Adaptive tests that adjust difficulty based on student responses create a more personalized and fair evaluation system. Additionally, integrating interactive simulations and virtual labs can offer hands-on learning experiences and real-time skill assessments.

Addressing ethical and bias concerns is crucial for maintaining fairness and transparency in AI-driven assessments. Using diverse and representative datasets for training AI models and conducting regular audits can help minimize bias. Ensuring AI systems are explainable allows educators and students to understand how assessment decisions are made. Fostering a collaborative learning environment further enhances the educational experience. AI can facilitate peer review processes by providing structured guidelines and rubrics for constructive feedback, encouraging students to engage in meaningful discussions about AI use in education.

Continuous monitoring and improvement are necessary for the successful implementation of AI in education. Establishing feedback mechanisms where students and educators can report concerns helps refine AI tools and assessment methods. Organizing workshops where students and educators collaboratively design assessment tasks, brainstorm ideas, and engage in hands-on activities can enhance AI-driven assessments. Regularly seeking student feedback and using it to refine assessments ensures their ongoing relevance and effectiveness. Involving students in the creation of rubrics and assessment criteria helps them understand expectations and promotes fairness and transparency. Implementing peer review processes where students evaluate each other's work using AI-generated guidelines fosters collaborative learning and critical thinking. Designing assessments that require students to solve real-world problems using AI tools makes learning more engaging and applicable to practical situations. Educating students on the ethical implications of AI in assessments and involving them in discussions on academic integrity and responsible AI use further promotes a balanced and informed approach to AI integration in education.

Conclusion

Generative Artificial Intelligence holds significant potential to revolutionize classroom assessments by providing personalized, adaptive, and interactive evaluation methods tailored to students' diverse needs. AI tools can automate grading, create customized learning materials, and facilitate peer assessments, thereby enhancing both formative and summative evaluations. However, the adoption of AI in education requires careful consideration of ethical issues, such as maintaining academic integrity, addressing bias in AI-generated content, and ensuring equitable access to technology. Additionally, involving students in the

design and implementation of AI-driven assessments is crucial to promoting transparency, fairness, and a collaborative learning environment. As educational technologies progress, strategically using generative artificial intelligence can improve the quality and effectiveness of assessments, ultimately leading to better educational outcomes for students.

Suggestions

- Provide ongoing AI literacy training for educators on the ethical use of Generative AI in assessments.

- Develop clear ethical policies on AI use in assessments, focusing on privacy, bias, transparency, and integrity.

- Use AI to create adaptive, personalized assessments tailored to individual student performance.

- Ensure equitable access to AI resources for all students by addressing the digital divide.

References

Afemikhe, O. A. (2019). *Innovations in learning-oriented assessment practices.* Keynote address presented at the AEAA conference, Abuja.

Bala, K., Colvin, A., Christiansen, M. H., Heinemann, A. W., Kreps, S., Levine, L., … & Ziewitz, M. (2023). *Generative artificial intelligence for education and pedagogy.* Cornell University. Retrieved from https://arxiv.org/abs/2303.00310

Beckingham, S., Lawrence, J., Powell, M., & Hartley, J. (2024). *Using generative AI effectively in higher education: Sustainable and ethical practices for learning, teaching, and assessment.* Routledge.

Binns, T. (2022). Addressing bias in AI-driven assessments. *Journal of Artificial Intelligence in Education, 35*(2), 14-28.

BMC Medical Education. (2024). Artificial intelligence and medical education: Application in classroom assessment. *BMC Medical Education, 24*(1), Article 45. https://doi.org/10.1186/s12909-024-03567-7

Budiyono, B., & Mardiyana, M. (2019). Assessment for learning in mathematics education: An analysis of its implementation and challenges. *Journal of Mathematics Education, 10*(1), 45-52. https://doi.org/10.22342/jme.10.1.6789

Clos, J., & Chen, Y. Y. (2024). Investigating the impact of generative AI on students and educators: Evidence and insights from the literature. *Proceedings of the*

Second International Symposium on Trustworthy Autonomous Systems. https://doi.org/10.1145/3565648

Cohen, L., Manion, L., & Morrison, K. (2020). Assessment for learning: Current practice and future directions in educational contexts. In *Research methods in education* (8th ed., pp. 399-412). Routledge. https://doi.org/10.4324/9781315456539

Cornell University. (2023). *Generative artificial intelligence for education and pedagogy.* Cornell University. Retrieved from https://arxiv.org/abs/2303.00310

Dwivedi, Y. K., Hughes, D. L., Kar, A. K., Baabdullah, A. M., Grover, P., Abbas, R., & Sharma, S. K. (2023). Ethical AI for teaching and learning. *Journal of Business Research, 145,* 456-467. https://doi.org/10.1016/j.jbusres.2022.05.042

Hadley, J. A., & Ardito, C. G. (2024). Generative AI in assessment: Towards understanding the student view. *MSOR Connections.* https://doi.org/10.21100/msor.v24i1.123

Hendri, S., Tjendami, I., & Hamid, A. (2019). Assessment for learning: Enhancing student engagement and learning outcomes in higher education. *International Journal of Educational Research, 95,* 89-98. https://doi.org/10.1016/j.ijer.2019.01.002

Hidayat, M., Rosmaiyadi, R., & Zubaidah, Z. (2023). Assessment for learning: Implementation and its impact on students' academic achievements. *Journal of Education and Practice, 14*(3), 55-63. https://doi.org/10.7176/JEP-14-3-05

IBM. (2024). *What is generative AI?* Retrieved from https://www.ibm.com/topics/generative-ai

Imasuen, K., & Adeosun, K. (2022). Assessment: Implication for national development. *Unizik Journal of Educational Research and Policy Studies, 14*(2), 53-64.

Miller, P. (2022). Human judgment in AI-driven assessments: Challenges and solutions. *Educational Technology Research and Development, 70*(1), 45-58.

MIT News. (2023). The rise of generative AI in education. *MIT News.* Retrieved from https://news.mit.edu/2023/rise-generative-ai-education

Perkins, M., Furze, L., Roe, J., & MacVaugh, J. (2024). The AI assessment scale (AIAS): A framework for ethical integration of generative AI in educational assessment. *Journal of University Teaching & Learning Practice, 21*(2), 123-145. https://doi.org/10.53761/q3azde36

Russell, S. J., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson Education Limited.

Scheelenkans, E., Vermeulen, G., Tijms, J., van der Kooij, D., & van der Linden, W. (2021). Assessment for learning: The impact of formative assessment on student learning in higher education. *Assessment & Evaluation in Higher Education, 46*(6), 935-950. https://doi.org/10.1080/02602938.2021.1907977

Tjendami, I., Hendri, S., & Hamid, A. (2019). Exploring effective assessment practices for learning in higher education. *Journal of Learning and Teaching in Higher Education, 18*(3), 203-218. https://doi.org/10.1080/13562517.2019.1571324

VanLehn, K. (2023). Teacher readiness for AI tools in education. *Journal of Educational Technology, 62*(3), 75-90.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems, 30*, 5998-6008.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education, 16*(1), 1-27.

Zhai, X. (2022). Artificial intelligence in formative assessment: Opportunities and challenges. *Computers & Education: Artificial Intelligence, 3*, 100054.