NEW ERA INTERNATIONAL JOURNAL OF INTERDISCIPLINARY SOCIAL RESEARCHES ISSN 2757-5608

Article Arrival Date
05. 06. 2025

Article Type
Research Article

Article Published Date
20.10.2025

Doi Number: 10.5281/zenodo.17391568

A REVIEW OF THE USE OF GENERATIVE ARTIFICIAL INTELLIGENCE IN ONLINE ASSESSMENT PROCESSES

Emin ÖZEN
Dr. (Ph.D.), Türkiye Republic Ministry of National Education, Assesment and Evaluation Department, Eskişehir, Türkiye
e-mail: eminozen@anadolu.edu.tr; ORCID ID: 0000-0001-7026-1503

Abstract

With the widespread use of online education environments today, the effectiveness and reliability of assessment processes have become increasingly important. In this context, artificial intelligence (AI) technologies are transforming online assessment processes through various applications such as automatic question generation, response evaluation, personalised feedback, and enhanced exam security. This paper examines the areas of application, advantages, and limitations of AI in online assessment processes. AI-based systems can generate questions in various formats by analysing educational materials, perform automatic evaluation of open-ended responses, and provide personalised feedback to students to help them correct their mistakes. Additionally, they can detect suspicious behaviour during exams to identify cheating attempts, thereby ensuring academic integrity. One of the most important advantages of these technologies is that they save educators time while making the assessment process more objective and reliable. However, there are some limitations, such as the accuracy, impartiality, and ethical issues of the responses generated by AI. In addition, the effectiveness of personalised feedback and the transparency of AI decision-making mechanisms are among the issues that need to be addressed. In conclusion, the opportunities offered by AI in online assessment processes bring about a significant transformation in education. However, interdisciplinary research and more comprehensive studies are needed to ensure that these technologies are used in a more reliable and ethical manner.

Keywords: Generative Artificial Intelligence, Online Measurement and Assessment, Automatic Assessment, Personalised Feedback, Exam Security, Educational Technologies, Online Distance Learning

Introduction

In recent years, digital transformation has led to profound and comprehensive changes in the field of education, radically altering the structure of teaching and learning processes (Selwyn, 2016). In particular, the extraordinary circumstances experienced globally during the COVID-19 pandemic have made it necessary for education systems to adapt quickly to digital platforms (Dhawan, 2020). This situation has resulted in the widespread adoption of online education models and the standardisation of distance learning. As a result, the importance of digital solutions in measurement and evaluation (M&E) applications integrated into educational processes has increased, and various challenges have emerged in adapting traditional measurement tools to digital environments (Reeves & Lin, 2020).

When traditional assessment methods are used in online environments, fundamental measurement principles such as validity and reliability are threatened; factors such as technical infrastructure deficiencies, candidate identity verification issues, and objective assessment difficulties arise (Bennett, 2015). In this context, the use of technological innovations has become inevitable to enhance the effectiveness of measurement processes and improve the quality of assessment. Artificial intelligence (AI)-based technologies, in particular, are among the most critical tools in this transformation in education, providing capabilities such as automation, personalisation, and data analytics in measurement and assessment processes (Heffernan & Heffernan, 2014).

Among these technologies, generative artificial intelligence (GAI) systems, which are gaining increasing attention in the field of education, demonstrate high performance in tasks such as text generation, content analysis, and question and answer creation thanks to developments in natural language processing (NLP) (Brown et al., 2020). Large language models (LLMs) such as ChatGPT, Claude, Gemini, and LLaMA from the GPT (Generative Pre-trained Transformer) family are not limited to chat and text generation but are also effectively used in areas such as the automation of measurement and evaluation processes in education, the analysis of student responses, and the provision of feedback to teachers (Bommasani et al., 2021). The advanced capabilities of AI have the potential to reduce human error in the assessment process, thereby objectifying evaluation, and can significantly reduce teachers' workload (Heffernan & Heffernan, 2014).

However, AI-based measurement and assessment applications also have important limitations, such as ethics, privacy, algorithmic bias, and measurement fairness (Bender et al., 2021). Therefore, it is necessary to comprehensively examine these applications of AI in education and develop guiding principles and standards.

The primary objective of this review study is to present the areas of application, advantages, and limitations of AI in online assessment and evaluation processes within an academic and conceptual framework. Additionally, by discussing research gaps and potential future research directions regarding the ethical and effective use of AI in education, the study aims to provide a comprehensive perspective for educators, researchers, and policymakers.

Theoretical Foundations and Definition of Generative Artificial Intelligence

Assessment and evaluation processes play a critical role in determining whether educational objectives have been achieved and in identifying learners' levels of knowledge, skills, and competencies. In Bloom's classification of learning domains, assessment is considered the highest level of cognitive development, while in Biggs and Tang's (2007) constructivist learning theory, the assessment process is defined as part of learning. In this sense, measurement and assessment are components that not only evaluate results but also nourish and guide the learning process. With the transfer of these processes to digital environments, the inadequacies of classical approaches have become more apparent, and the need for technological tools has rapidly increased.

Generative artificial intelligence (GAI) stands out as a powerful technological solution that can address this need, particularly in the field of education. GAI systems are typically trained on large datasets and utilise statistical learning methods and deep learning algorithms. These systems not only analyse existing data but also have the capacity to generate original content (Brown et al., 2020). Large language models (LLMs) such as ChatGPT, Gemini, and Claude are among the most common representatives of generative artificial intelligence. These models can generate original questions, explanations, summaries, and evaluation criteria using natural language processing (NLP) techniques.

In an educational context, AI enables multi-dimensional applications, particularly in online measurement and evaluation processes, such as automatic question generation, analysis of students' responses, provision of feedback, and ensuring exam security. For example, tasks such as generating meaningful and taxonomy-appropriate questions from a text passage, classifying open-ended responses based on their level of understanding, or tailoring feedback according to students' error tendencies—which are limited by traditional algorithms—can be executed

flexibly and contextually by AI systems (Pan et al., 2022). In this sense, generative artificial intelligence is not merely a technical tool; it also serves as a learning assistant capable of performing pedagogical functions.

However, technical proficiency alone is insufficient for the use of GAI in educational assessment processes. There is also a need for designs that serve pedagogical purposes, are sensitive to ethical principles, and support human-AI interaction (Holmes et al., 2019). Therefore, the effective and sustainable use of AI applications in education requires the contribution of not only artificial intelligence but also educational sciences. In the following sections of this review study, the concrete applications of AI in the field of online measurement and evaluation will be detailed within this framework.

Generative Artificial Intelligence-Based Measurement and Assessment Applications

Automatic Question and Assessment Tool Production

Generative artificial intelligence (GAI) technologies are fundamentally transforming the question creation process in education. In traditional methods, teachers individually prepare questions appropriate to the curriculum content and conduct meticulous work to ensure that these questions are consistent with a comprehensive taxonomic structure. GAI, on the other hand, automates this process through large data sets and deep learning algorithms. Systems can generate questions based on learning objectives and pedagogical principles, referencing theoretical frameworks such as Bloom's cognitive taxonomy. This enables the creation of rich question pools that include both lower cognitive levels (knowledge, comprehension) and higher cognitive levels (analysis, synthesis, evaluation) (Pan et al., 2022). Additionally, these systems increase assessment diversity by effectively generating questions in different formats—multiple choice, true/false, matching, open-ended. Especially in environments where large groups of students are assessed, this automation significantly reduces teachers' workload while standardizing and improving the quality of assessment tools (Kellogg et al., 2021). Furthermore, thanks to the adaptive structure of AI, the difficulty level of questions can be dynamically adjusted according to students' different learning speeds and levels, enabling individualized measurement and assessment processes.

The natural language processing capabilities of generative artificial intelligence enable in-depth analyses beyond traditional systems, particularly in the evaluation of open-ended questions. In this context, the following features of generative artificial intelligence are noteworthy:

- Semantic and Contextual Analysis: AI can evaluate conceptual consistency, originality, cause-and-effect relationships, and critical thinking skills in students' written responses (Bennett, 2015; Zhai & Kong, 2022). Thus, evaluation is based on meaning rather than just words.

- Consistent and Objective Assessment: Biases and scoring inconsistencies seen in human assessment are minimized with AI. It provides systematic and objective scoring, especially in large student groups (Zhai & Kong, 2022).

- Instant Feedback: Students can receive quick and detailed feedback after submitting their responses. This feedback supports the learning process in real time and increases motivation (Holmes et al., 2019).

• Evaluation of Complex Responses: AI systems can even assist in the evaluation of complex written compositions, essays, and project reports (Shermis & Burstein, 2013).

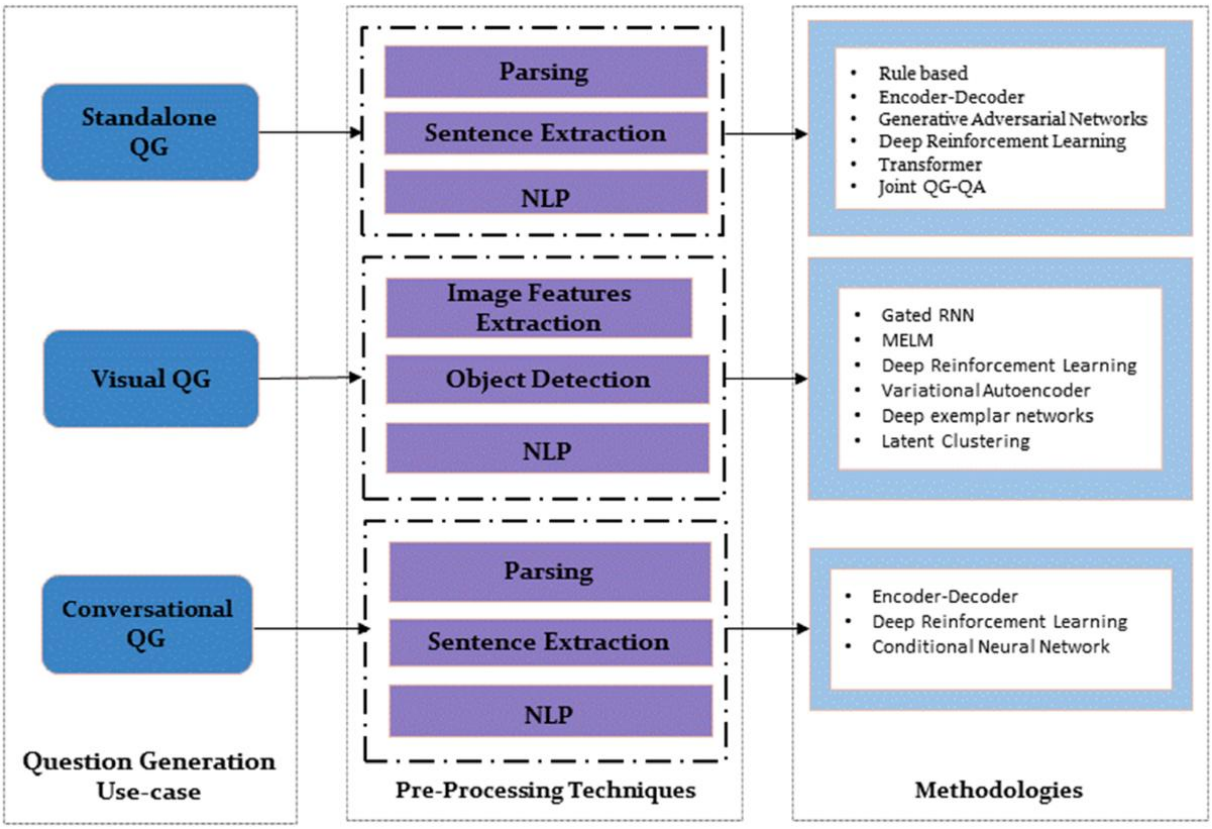

Fig.1. Automatic question generation (Mulla & Gharpure, 2023)

Automatic Response Analysis and Scoring Systems

Open-ended questions are particularly important in education because they reveal students' deep thinking and expression skills, but evaluating such responses is time-consuming and subjective. Generative artificial intelligence analyzes students' responses at the semantic level through natural language processing techniques. This analysis includes not only whether the answers are correct or incorrect, but also parameters such as the conceptual consistency, originality, logical connections, level of critical thinking, and suitability for pedagogical purposes of the answers (Zhai & Kong, 2022). Thus, the scoring process becomes more objective, free from biases arising from the human factor. Additionally, these systems quickly process hundreds or even thousands of responses, reducing the evaluation burden on teachers and increasing the efficiency of the evaluation process. More importantly, they provide instant feedback on the strengths and weaknesses of responses in a way that supports students' learning processes. This enables teachers to intervene more consciously and increases students' awareness of their own learning processes (Holmes et al., 2019).

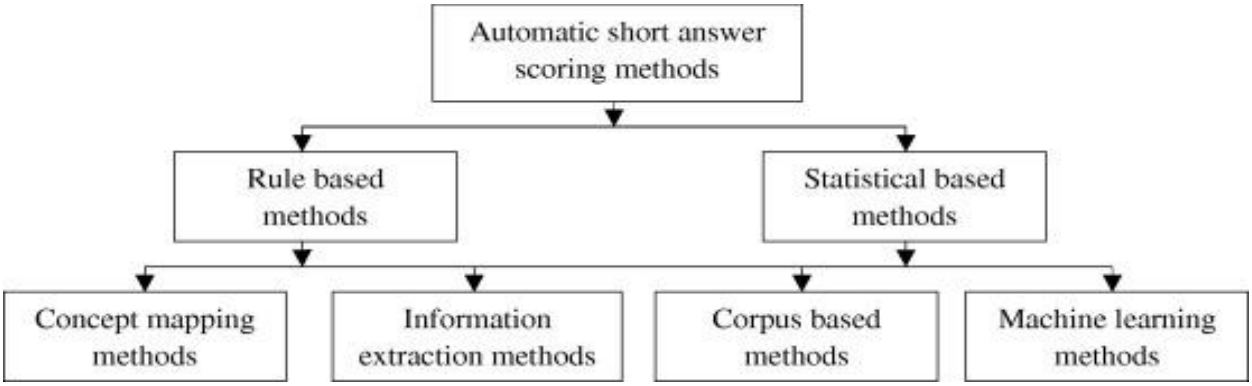

Fig. 2. Methods of Automatic short answer scoring (Artz et all., 2008)

Personalized Feedback and Learning Support Systems

Feedback is crucial for increasing students' active participation in the learning process. Productive AI-based systems not only provide information on which questions were answered incorrectly by analyzing student responses in detail, but also identify the causes of students' learning deficiencies and provide personalized, pedagogically meaningful feedback tailored to these deficiencies (Luckin et al., 2016). This feedback can be adapted to suit students' cognitive levels, learning styles, and motivations; for example, it can be enriched with visual, auditory, or textual support. In this way, students have the opportunity to deepen their learning processes rather than focusing solely on results. In addition, the adaptive learning paths provided by AI contribute to the development of students' own learning speeds and strategies. This approach plays a critical role, particularly in the development of self-regulation skills, and lays the foundation for lifelong learning habits by increasing learners' independence (Kellogg et al., 2021). This in-depth and personalized feedback structure increases students' motivation while also increasing learning retention and success.

Generative artificial intelligence analyzes student performance in detail and provides customized feedback tailored to individual learning needs, supporting the learning process. In this context, the following features of generative artificial intelligence are noteworthy:

- Error Detection and Guidance: Common errors made by students in their responses are identified, and explanations, examples, and additional materials are suggested to address these deficiencies (Kellogg et al., 2021). This is an effective method for closing the learning gap.
- Feedback Tailored to Learning Style: Feedback can be provided in different formats (written, visual, audio) according to students' cognitive and affective characteristics (Luckin et al., 2016).
- Self-regulation and autonomy: Feedback processes guide students in monitoring and regulating their own learning, thereby supporting learner autonomy (Holmes et al., 2019).
- Adaptive Learning Pathways: Personalizes the learning experience by providing students with different learning materials and exercises based on their performance (Heffernan & Heffernan, 2014).

Exam Security and Ethical Monitoring Systems

With the widespread use of online exams, ensuring exam security and academic integrity has presented new challenges. Generative artificial intelligence technologies can perform behavioral analysis by integrating data from multiple sources (camera images, microphone recordings, keyboard movements,

browser activities) and detect unusual exam behavior in real time (Zhou et al., 2023). These systems contribute significantly to the fairness and security of the exam process by accurately identifying ethical violations such as cheating, receiving outside help, or failing to comply with exam rules. However, these advanced surveillance techniques also raise serious ethical concerns, particularly regarding student privacy and data security. Full compliance with national and international data protection regulations is required during the collection, storage, and processing of student data (Floridi et al., 2018). Additionally, it is critical that AI decision-making mechanisms are not a "black box"; they must be transparent, explainable, and open to human oversight to prevent potential misjudgments. In this context, ethical principles should be an integral part of educational institutions' exam security policies.

The security of online exams is ensured by advanced surveillance and anomaly detection systems supported by AI technologies. In this context, the following features of generative artificial intelligence are noteworthy:

• Real-Time Behavioral Monitoring: By analyzing multiple data sources such as camera, microphone, and keyboard movements, unusual and suspicious exam behaviors are detected (Zhou et al., 2023).

• Anomaly Algorithms and Artificial Intelligence: Attempts to cheat, seek help from others, or access external sources during exams are detected with high accuracy using machine learning algorithms (Kellogg et al., 2021).

• Data Privacy and Ethical Standards: Data should be collected and processed while protecting student privacy. Transparent and fair practices that comply with international data protection laws should be developed (Floridi et al., 2018).

• Importance of Human Oversight: AI-supported surveillance systems should be supported by human oversight and decision-making processes to prevent automated decisions from leading to erroneous or unfair outcomes (Burrell, 2016).

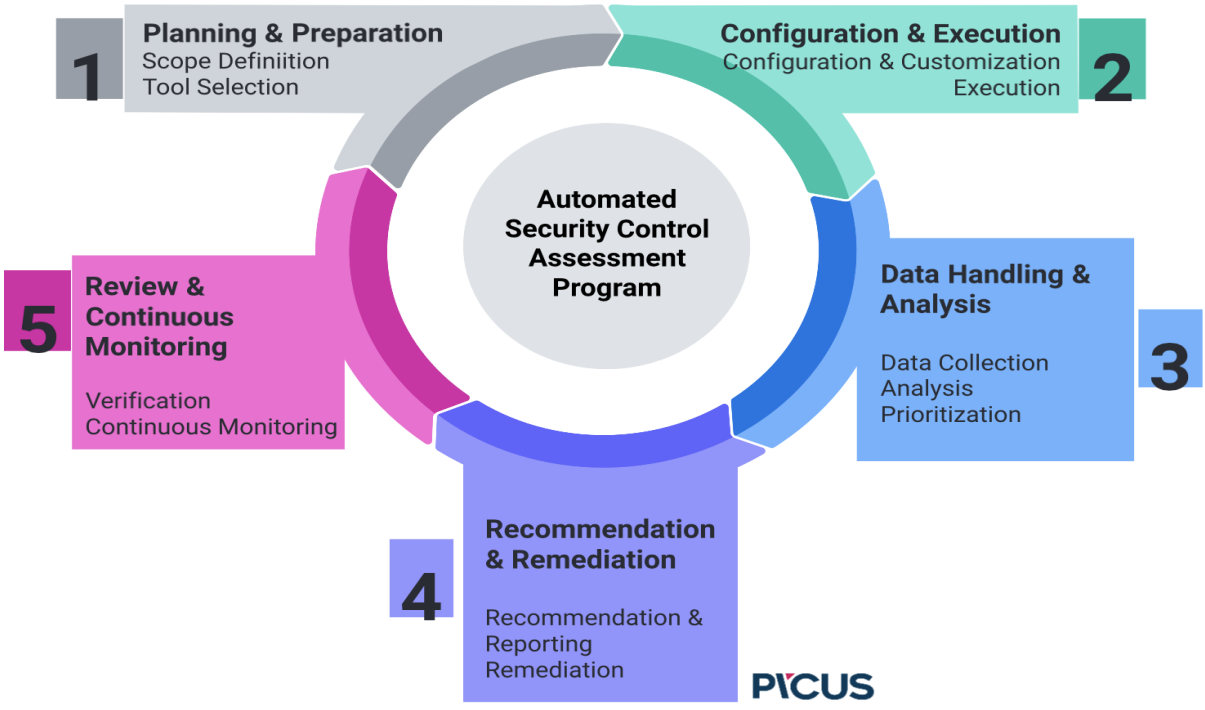

Figure 3. Life-cycle of an Automated Security Control Assessment Program**.**

Learning Analytics and Decision Support Systems

Generative artificial intelligence analyzes large data sets obtained from educational processes, offering new opportunities in the field of learning analytics. These systems monitor and interpret students' individual performance trends, learning difficulties, motivation levels, and behavioral patterns in a multidimensional manner (Luckin et al., 2016). The data obtained provides scientific-based support to teachers, students, and education administrators in strategic decision-making processes. Thus, effective interventions can be designed not only for individual students but also based on performance data at the class, school, and education system levels. For example, targeted support programs can be created for student groups with poor performance, and teaching methods can be restructured according to this data. In addition, data-driven and objective approaches are becoming more widespread in the formulation of education policies (Williamson & Eynon, 2020). Generative AI-based learning analytics not only improve quality and equity in education but also strengthen teachers' pedagogical decisions and contribute to the sustainability of student achievement.

Thanks to its big data analysis capabilities, generative artificial intelligence evaluates student performance in a multidimensional way and provides support to education stakeholders for strategic decisions. In this context, the following features of generative artificial intelligence are noteworthy**:**

• Real-Time Performance Monitoring: Students' achievement levels, areas of difficulty, and rates of progress are tracked in real time (Pan et al., 2022).

• Data-Driven Intervention Plans: Teachers and administrators can develop targeted interventions and teaching strategies based on analyzed data (Luckin et al., 2016).

• Scientific Foundations of Education Policies: Systematic data analysis provides a scientific basis for the design and implementation of education policies (Williamson & Eynon, 2020).

• Analysis at the Student, Classroom, and Institutional Levels: ÜYZ systems analyze not only individual but also collective performance data, offering opportunities for improvement across all components of the education system (Floridi et al., 2018).

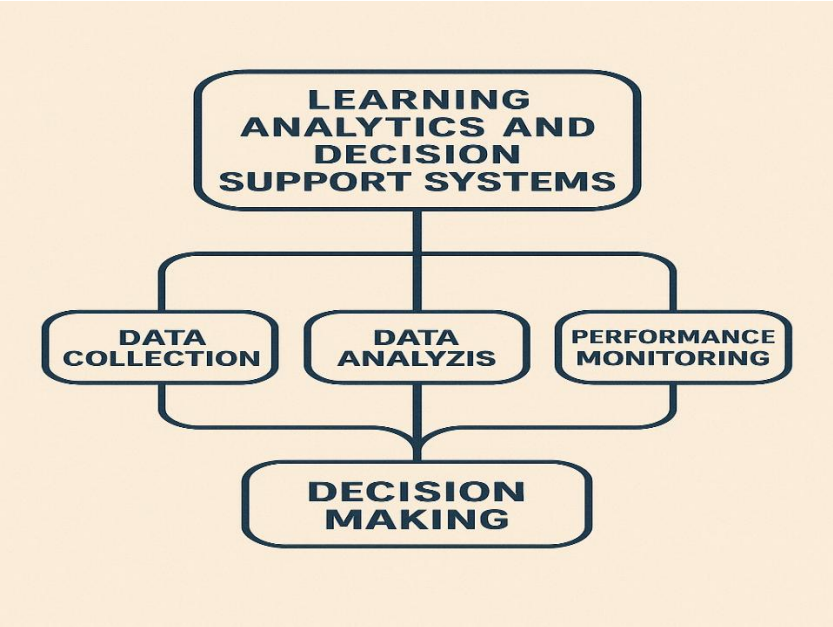

Fig.4. Learning Analytics and Decision Support System Components (Dall-E, 2025)

Advantages and Opportunities

The integration of generative artificial intelligence (GAI) into online measurement and assessment processes offers significant opportunities in education, both pedagogically and operationally. Chief among these opportunities is the saving of time and labor. Preparing measurement tools, evaluating student responses, and providing feedback using traditional methods is a time-consuming and labor-intensive process. AI systems significantly automate these processes, reducing teachers' workload and enabling them to focus on more creative, guiding, and mentoring tasks (Kellogg et al., 2021). This is particularly advantageous for teachers with large student groups.

Another important contribution is that it increases the objectivity and consistency of the assessment process. Human assessment can often vary depending on subjective judgments, cognitive biases, or assessment fatigue. However, AI systems produce more consistent and impartial results because they assess based on predefined criteria and algorithms. Especially in the evaluation of open-ended responses, AI can provide consistent scoring by systematically analyzing the responses of a large number of students (Zhai & Kong, 2022). In this respect, it has the potential to support equal opportunities in education.

In addition, the personalization opportunities offered by AI systems are largely consistent with the student-centered learning approach. Each student's learning speed, knowledge level, and tendency to make mistakes are different. Generative artificial intelligence can provide individualized feedback by analyzing student responses and offer support in areas where the student is lacking (Holmes et al., 2019). This supports students in becoming aware of their own learning processes and developing more effective learning strategies. In this way, assessment ceases to be merely a tool for determining results and becomes a process that deepens learning.

AI also offers various opportunities in terms of inclusive education. Providing adaptable content for students with different languages, cultures, learning levels, or special needs is quite challenging with traditional methods. However, AI can cater to different learner profiles with features such as multilingual support, adaptable difficulty levels, and visual/auditory explanations. This increases the accessibility of learning environments and enables disadvantaged groups to participate more effectively in education.

Finally, learning analytics and data-driven decision-making processes are among the important opportunities offered by AI. AI systems can monitor student performance in real time, obtain meaningful data, and provide decision support systems to teachers, students, and administrators based on this data. As a result, assessment results provide more qualitative feedback not only at the individual level but also at the class, school, and system levels (Luckin et al., 2016). Such data-driven approaches are gaining increasing importance in shaping education policies.

Despite all these advantages, the adoption of AI applications is not possible in an unlimited and unconditional manner. There are certain limitations, risks, and ethical responsibilities associated with the use of these technologies. The next section addresses these limitations in greater detail.

Limitations and Ethical Debates

Although generative artificial intelligence (GAI) technologies offer great potential for online measurement and assessment processes, the unrestricted and unconditional adoption of these

technologies carries serious risks. Therefore, the use of GAI-based systems in education must be addressed from both technical and ethical perspectives; regulations must be developed to ensure that these systems are reliable, fair, and accountable. Otherwise, rather than the technology being used for good purposes, systematic errors and inequalities may be reproduced (Williamson & Eynon, 2020).

First, the accuracy and reliability of the content produced by AI is still up for debate. Even though these systems are super successful at generating language, they can sometimes produce semantic errors, contextual deviations, and content that isn't pedagogically appropriate. This can cause serious problems, especially when students are guided and evaluated using this content. In addition, the reflection of biases in the data used to train some models in the results produced is also a significant problem (Bender et al., 2021). For example, situations in which students may be discriminated against or systematically receive lower scores in socio-cultural content should not be overlooked.

Secondly, the lack of transparency in AI decision-making processes is another important limitation that can undermine educational reliability. The “black box” nature of artificial intelligence systems often prevents users from knowing how the system works, what criteria it uses for evaluation, or what data it uses to reach its conclusions (Burrell, 2016). This situation can deprive students and teachers of their right to question evaluation results and may even cause problems in judicial processes. In an area such as education, which directly affects an individual's future, the transparency and explainability of algorithmic calculations are of great importance.

Thirdly, students' excessive dependence on generative artificial intelligence can negatively affect their learning motivation and cognitive development. Students who constantly receive correct answers, ready-made explanations, or solutions from artificial intelligence may eventually stop using their own thinking and problem-solving skills. This situation may increase learning laziness and superficial learning tendencies in the long term (Selwyn, 2019). In addition, false successes may emerge that could lead to the performance of artificial intelligence rather than that of the student being measured in the assessment process.

An important aspect of ethical debates also concerns data privacy and student confidentiality. AI systems continuously analyze and process students' assessment responses, behaviors, and learning histories. Who stores such personal data, for what purpose, for how long, and with whom it is shared is often unclear. This situation increases students' digital traceability and may lead to violations of national and international data protection laws (Floridi et al., 2018). The interaction of children and young people with these systems requires a more rigorous regulatory and oversight process.

Finally, the lack of equipment for teachers and students to use artificial intelligence systems correctly and ethically limits the potential of technology. Artificial intelligence literacy is not just a technical competence; it is a multidimensional skill that requires critical digital thinking, ethical awareness, and pedagogical intuition. In this context, teachers should be encouraged to use AI-based assessment tools not only as implementers but also as questioning, developing, and guiding actors (Holmes et al., 2019).

All these limitations point to the need for technology to be structured in a more reliable, fair, and ethical manner, rather than constituting a rejection of the use of generative artificial intelligence in education. Therefore, it is of great importance that multidisciplinary research

teams, policymakers, and educators work together in the development and implementation of these systems**.**

Conclusions and Recommendations

This review study has comprehensively revealed the multidimensional effects of integrating generative artificial intelligence (GAI) into online measurement and assessment processes. The applications of AIG technologies in critical areas such as automatic question generation, open-ended response assessment, personalized feedback provision, and exam security hold significant potential in terms of quality, accessibility, and inclusivity in education (Brown et al., 2020; Holmes, Bialik & Fadel, 2019). In particular, in automatic question generation processes, AI reduces teachers' workload by designing questions that are aligned with learning objectives and include taxonomic diversity, while expanding the scope of assessment. Similarly, AI systems that use natural language processing techniques can provide objective, consistent, and pedagogically rich feedback through in-depth analysis of open-ended responses (Zhai & Kong, 2022). This effectively supports learners' learning processes and makes teachers' assessment processes more effective and efficient.

However, the widespread adoption of these technologies and ensuring their ethical use also present significant challenges and risks. Firstly, the accuracy and pedagogical appropriateness of the content produced by AI systems is not yet fully guaranteed; it can sometimes lead to semantic deviations and biased results (Bender, Gebru, McMillan-Major & Shmitchell, 2021). This situation can lead to students being misguided by incorrect information or unfair assessment. Additionally, the “black box” nature of AI in decision-making processes leads to a lack of algorithmic transparency and makes it difficult for users to understand and monitor the evaluation processes (Burrell, 2016). In a sensitive field such as education, which directly affects an individual's future, these transparency issues open the door to ethical and legal problems. Furthermore, excessive reliance on AI systems can negatively affect students' cognitive development and critical thinking skills (Selwyn, 2019). Therefore, artificial intelligence should be positioned as a complementary tool in learning processes.

Data privacy and confidentiality are critical ethical issues associated with the use of AI in education. The processes of collecting, processing, and storing student data must be transparent, secure, and compliant with regulations (Floridi, Cowls, Beltrametti, Chatila, Chazerand, Dignum et al., 2018). Especially in educational environments involving children and young people, data protection standards must be applied at the highest level and continuously monitored. In this context, it is essential that national and international regulatory bodies take a proactive role and that ethical standards are adopted by educational technology developers (Williamson & Eynon, 2020).

The sustainable and effective use of generative artificial intelligence in education requires multidisciplinary and participatory approaches that encompass not only technological developments but also pedagogical, ethical, and social dimensions. In this context, interdisciplinary collaborations between education scientists, artificial intelligence engineers, ethics experts, lawyers, and sociologists should be increased (Bommasani, Hudson, Adeli, Altman, Arora et al., 2021). Such collaborations will enable the development and implementation of AI systems in accordance with ethical principles, from design to application. In addition, it is imperative to develop comprehensive education programs on artificial intelligence literacy so that teachers and students can critically use artificial intelligence from pedagogical and ethical perspectives (Holmes et al., 2019). These trainings will not only

increase users' trust in technology but also contribute to the establishment of a culture of conscious and ethical use.

It is also of great importance to monitor and evaluate the effects of generative artificial intelligence applications on education in the long term using scientific and empirical methods (Luckin, Holmes, Griffiths & Forcier, 2016). In this context, the effects of AI on learning performance, motivation, teacher attitudes, and educational equity should be systematically researched, and policies and practices should be developed based on the data obtained. In particular, inclusive design principles should be adopted to prevent the deepening of digital inequalities; adaptable and accessible assessment tools should be developed and disseminated for students from different cultures, language groups, and learning needs (Holmes et al., 2019)..

In conclusion, while generative artificial intelligence has the potential to fundamentally transform measurement and assessment in education, multi-dimensional, conscious, and ethical approaches are essential for these technologies to be converted into educational benefits. The advancement of technological developments in tandem with pedagogical vision and ethical responsibility is key to the sustainable and equitable use of AI in education. Therefore, ongoing dialogue and collaboration between policymakers, educational institutions, and technology developers is expected to be decisive in shaping the educational environments of the future.

References

Arzt, G., Widmer, S., & Dixon, S. (2008). Automatic page turning for musicians via real-time machine listening. In *Proceedings of the 18th European Conference on Artificial Intelligence (ECAI 2008)* (pp. 241–245).

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623). doi:10.1145/3442188.3445922

Bennett, R. E. (2015). The changing nature of educational assessment. *Review of Research in Education, 39*(1), 370–407. doi:10.3102/0091732X14554179

Biggs, J., & Tang, C. (2007). *Teaching for quality learning at university* (3rd ed.). McGraw-Hill.

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., ... Liang, P. (2021). On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258*. Retrieved from http://arxiv.org/abs/2108.07258

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems, 33*, 1877–1901.

Burrell, J. (2016). How the machine ‘thinks’: Understanding opacity in machine learning algorithms. *Big Data & Society, 3*(1), 1–12. doi:10.1177/2053951715622512

Dhawan, S. (2020). Online learning: A panacea in the time of COVID-19 crisis. *Journal of Educational Technology Systems, 49*(1), 5–22. doi:10.1177/0047239520934018

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines, 28*(4), 689–707. doi:10.1007/s11023-018-9482-5

Gikandi, J. W., Morrow, D., & Davis, N. E. (2011). Online formative assessment in higher education: A review of the literature. *Computers & Education, 57*(4), 2333–2351. doi:10.1016/j.compedu.2011.06.004

Heffernan, N. T., & Heffernan, C. L. (2014). The ASSISTments ecosystem: Building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching. *International Journal of Artificial Intelligence in Education, 24*(4), 470–497. doi:10.1007/s40593-014-0024-x

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.

Kellogg, S., Booth, S., & Oliver, K. (2021). A social network perspective on peer supported learning in MOOCs for educators. *International Review of Research in Open and Distributed Learning, 22*(2), 243–263. doi:10.19173/irrodl.v22i2.5239

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson.

OpenAI. (2023). *GPT-4 technical report*. *arXiv preprint arXiv:2303.08774*. Retrieved from http://arxiv.org/abs/2303.08774

Pan, L., Liu, J., & Li, L. (2022). Automatic question generation for educational applications: A systematic review. *IEEE Transactions on Learning Technologies, 15*(2), 226–244. doi:10.1109/TLT.2022.3157063

Reeves, T. C., & Lin, L. (2020). The research we have is not the research we need. *Educational Technology Research and Development, 68*(4), 1991–2001. doi:10.1007/s11423-020-09811-3

Selwyn, N. (2016). *Education and technology: Key issues and debates* (2nd ed.). Bloomsbury Academic.

Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education*. Polity Press.

Shermis, M. D., & Burstein, J. (Eds.). (2013). *Handbook of automated essay evaluation: Current applications and new directions*. Routledge.

Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology, 45*(3), 223–235. doi:10.1080/17439884.2020.1798995

Zhai, X., & Kong, X. (2022). A review of artificial intelligence in education from 2010 to 2020. *Journal of Educational Computing Research, 60*(2), 522–546. doi:10.1177/07356331211055377

Zhou, Y., Huang, C., Hu, Q., Zhu, J., & Tang, Y. (2023). Ensuring academic integrity in online assessments: A blockchain-based approach. *Computers & Education, 184*, 104516. doi:10.1016/j.compedu.2022.104516

.