



Artificial Intelligence in Classroom Assessment: Opportunities, Equity Challenges, and Best Practices for Formative and Summative Integration

Gilbert Kalonde, Samuel Boateng, Claudia Duedu

Department of Education, Montana State University, Bozeman, Montana, USA

Email: Kalonde@montana.edu

How to cite this paper: Kalonde, G., Boateng, S. and Duedu, C. (2025) Artificial Intelligence in Classroom Assessment: Opportunities, Equity Challenges, and Best Practices for Formative and Summative Integration. *Open Access Library Journal*, 12: e14121.

<https://doi.org/10.4236/oalib.1114121>

Received: August 15, 2025

Accepted: September 27, 2025

Published: September 30, 2025

Copyright © 2025 by author(s) and Open Access Library Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Artificial Intelligence (AI) is rapidly transforming classroom assessment by enabling adaptive, data-driven, and personalized approaches that extend beyond the limitations of traditional methods. This paper presents a systematic literature review of recent peer-reviewed studies focusing on AI integration in both formative and summative assessments. The review synthesizes evidence on four thematic areas: personalization, efficiency, and scalability; equity and inclusion challenges; ethical and transparency concerns; and methodological gaps in current research. Findings indicate that AI-powered tools—such as adaptive learning platforms, automated scoring systems, and natural language processing applications enhance feedback timeliness, assessment accuracy, and instructional responsiveness, particularly in high-resource settings with strong infrastructure, teacher training, and policy support. However, persistent challenges related to algorithmic bias, data privacy, transparency, and unequal access in low-resource contexts limit equitable adoption. Methodological imbalances in the literature, including overreliance on pilot studies in affluent contexts, further constrain generalizability. The paper concludes that realizing AI's transformative potential in classroom assessment requires equity-centered implementation, culturally responsive design, robust governance, and sustained professional development to ensure AI serves as a tool for inclusion rather than a mechanism for reinforcing educational disparities.

Subject Areas

Educational Technology, Assessment and Evaluation

Keywords

Artificial Intelligence, Classroom Assessment, Formative Assessment, Summative Assessment, Adaptive Learning, Automated Scoring, Educational Equity, Algorithmic Bias, Data Privacy, Educational Technology, Personalized Learning, Ethics in AI

1. Introduction

AI in computer science represents a significant milestone, defined by the capacity of digital systems such as computers or robots to perform tasks traditionally associated with human intelligence [1]. AI involves the simulation of human behavior and cognitive awareness, enabling machines to replicate intelligent actions [1]. It encompasses a wide range of capabilities, including cognitive automation, machine learning, reasoning, hypothesis testing, natural language processing, and adaptive algorithm modification, which together contribute to a deeper and often more efficient understanding of information [2].

The promise of AI has manifested in various applications such as recommendation systems, virtual assistants, facial recognition technologies, and educational tools [2]. Within K-12 education, algorithms and machine learning support personalized learning and efficient assessment, making AI a promising tool to enhance classroom engagement and improve student outcomes [3].

Recent developments illustrate the transformative potential of AI in classroom assessment. Traditional paper-and-pencil examinations and manual grading methods have proven inadequate in meeting the growing demand for personalization and timely feedback [4]. AI-driven systems introduce innovative approaches that increase both efficiency and effectiveness in assessment, including adaptive platforms, automated scoring tools, and real-time feedback mechanisms that individualize the assessment process [5].

AI technologies can process large datasets and provide personalized feedback to learners, enabling more informed and nuanced evaluation practices [6]. Tools such as ALEKS and DreamBox Learning, for example, use adaptive algorithms to adjust the difficulty of questions in real time based on student responses [6]. Natural language processing tools such as Grammarly offer immediate writing feedback that helps students develop their skills over time [7].

Despite these advantages, challenges remain. Concerns around data privacy, algorithmic bias, and the potential misuse of student information underscore the importance of equity and reliability in AI adoption [7]. Nevertheless, AI continues to grow in its application across educational institutions because of its demonstrated ability to enhance both formative and summative assessment practices [7]. While AI holds the potential to revolutionize classroom assessment, its effective integration requires deliberate strategies to address personalization, efficiency, ethics, and best practices. This study therefore investigates how AI can be effec-

tively integrated into classroom assessment, improve efficiency, overcome ethical challenges, and establish sustainable models for responsible implementation.

1.1. Problem Statement

AI integration in classroom assessments is rapidly expanding, yet significant gaps remain in understanding its full impact on teaching and learning. Traditional methods of assessment, largely dependent on static exams and manual grading, struggle to provide timely feedback or address students' individual learning needs [4] [8]. These limitations highlight the potential of AI to transform assessment practices through automation, personalization, and data-driven insights that enable more adaptive and responsive evaluation [1] [6] [9].

At the same time, the integration of AI into education raises concerns about data privacy, algorithmic bias, and the validity of AI-driven assessments. Scholars caution that without greater transparency and safeguards, AI systems risk reinforcing existing inequities rather than reducing them [7]. Furthermore, implementation barriers—such as resource constraints, lack of teacher training, and unequal access to technology, prevent many schools from effectively adopting AI-driven assessments [9].

This study addresses these challenges by critically reviewing existing research on AI's potential and limitations in both formative and summative assessments. It seeks to clarify how AI can improve efficiency and personalization while also ensuring that ethical and equity considerations remain central to its integration in diverse educational contexts.

1.2. Purpose of the Study

The purpose of this study is to explore the integration of AI in classroom assessments, with particular focus on how AI enhances personalized learning and feedback, improves assessment efficiency, addresses ethical challenges, and identifies best practices for effective implementation. By examining these dimensions, the study seeks to provide insights that support responsible and impactful AI adoption in educational assessment practices.

1.3. Research Questions

This study is guided by the following research questions:

- 1) How is AI integrated in formative and summative assessments?
- 2) What benefits does AI integration bring to assessing efficiency?
- 3) What ethical challenges accompany AI integration in assessments?
- 4) What best practices support effective AI integration in classroom assessments?

1.4. Significance

AI-powered assessments have the potential to transform education by enhancing efficiency, personalization, and data-driven practices. This study contributes to

the growing body of literature by providing an in-depth analysis of how AI tools can reshape teaching and learning through classroom assessment. It highlights the dual promise and challenges of AI, emphasizing not only its ability to save instructional time and provide immediate feedback but also the importance of addressing equity, transparency, and ethical use. The findings are intended to inform educators, policymakers, and researchers about how to implement AI responsibly, ensuring that innovation advances inclusion and educational effectiveness.

2. Literature Review

2.1. Integrating AI into Formative Assessments

Formative assessments are designed to monitor learners' ongoing progress and provide timely feedback that supports instructional adjustments in real time. Unlike summative evaluations that measure end results, formative assessments aim to inform the teaching process as it unfolds. In this regard, AI technologies have substantially enhanced formative practices by introducing adaptive digital tools that can analyze student performance continuously and provide individualized feedback without the delays and burdens associated with traditional manual assessments. By embedding AI-driven systems into the daily learning process, teachers can now identify gaps more precisely, intervene more promptly, and tailor instruction to meet diverse learner needs.

2.1.1. Adaptive Learning Systems

One of the most widely recognized contributions of AI in formative assessment is the use of adaptive learning systems. [10] emphasize that systems such as ALEKS and DreamBox Learning are designed to instantly adjust the difficulty of tasks based on student responses. This dynamic adjustment creates a personalized pathway for each learner, supporting mastery at an individual pace. Driven by AI algorithms, these platforms generate customized problem sets, explanations, and scaffolds that stimulate deeper engagement. Research confirms that learners in adaptive environments often outperform peers in traditionally structured classrooms, as they receive immediate, targeted support that reduces redundancy and accelerates progress [10]. Moreover, adaptive learning systems foster self-efficacy by creating self-paced environments that lower anxiety while maintaining high levels of motivation.

2.1.2. Natural Language Processing (NLP) Tools in Writing

Beyond numerical and conceptual domains, AI has also transformed writing instruction through Natural Language Processing (NLP). [11] argue that formative writing assessments benefit greatly from applications like Grammarly and WriteLab, which provide real-time feedback on grammar, coherence, and organization. These tools not only help students refine their writing while composing but also relieve instructors of repetitive error-checking tasks, thereby freeing instructional time. [12] further report that NLP-based assessment tools are espe-

cially effective in large-scale online courses, where providing frequent individualized feedback would be nearly impossible with traditional methods. Students who received AI-mediated feedback in such settings expressed greater satisfaction and demonstrated improved writing performance, attributing their progress to the immediacy and specificity of feedback compared to delayed, generalized comments from instructors [12].

2.1.3. Real-Time Feedback and Continuous Assessment

Another significant benefit of AI in formative assessment is the capacity to generate real-time feedback and continuous monitoring. [13] highlights that in conventional classrooms, students often wait days or weeks to receive grades, thereby missing critical opportunities to correct misconceptions. In contrast, AI-powered platforms such as Socrative and Classkick provide teachers with instant dashboards of student understanding, enabling immediate instructional adjustments. Real-time feedback has been shown to heighten engagement, as learners are motivated to correct errors on the spot [13]. For teachers, this reduces grading burdens and allows them to redirect their time to designing more engaging and higher-order tasks. As a result, classroom dynamics become more interactive and responsive, with both teachers and learners benefiting from faster cycles of evaluation and improvement.

2.1.4. Automated Scoring Systems

Perhaps the most promising and widely adopted use of AI in formative settings is automated scoring. These systems use algorithms to grade assignments, essays, and exams with high levels of efficiency and consistency [7]. Unlike traditional grading that is time-consuming and vulnerable to human error or bias, AI-based scoring ensures rapid feedback across large groups of students. For instance, major online course providers such as EdX and Coursera have relied on automated scoring to manage the overwhelming volume of submissions [9]. Likewise, more than 500 colleges have adopted Gradescope, a platform that highlights errors, underlines correct answers, and provides clear explanations to students. What distinguishes modern automated scoring systems from earlier versions is their ability to assess complex work like essays, not just multiple-choice questions [6]. These tools help students understand how to revise and improve, thus turning assessment into a learning experience rather than merely an evaluative one.

2.2. Integrating AI into Summative Assessments

Summative assessments evaluate student learning at the conclusion of an instructional unit, typically determining grades, certification, or progression. Incorporating AI into summative assessments has been revolutionary, introducing predictive analytics, machine learning, and automated grading at scales previously unimaginable. [14] argue that AI-driven summative assessments offer unprecedented reliability, speed, and flexibility by capturing nuanced data points and providing tailored insights. This section reviews the most salient AI applications

in summative assessment.

2.2.1. Automated Assessment and Evaluation Instruments

AI-driven platforms such as Gradescope and Turnitin have become highly influential in summative evaluation. [15] notes that Gradescope applies machine learning algorithms to classify and evaluate open-ended responses, ensuring consistency across large cohorts. This automation is particularly valuable for essay-style answers where maintaining fairness is often challenging. Similarly, Turnitin, initially known for plagiarism detection, has evolved to use AI for assessing originality and coherence in student work. [16] emphasize that such systems not only expedite grading but also reduce bias, enabling more equitable evaluations.

2.2.2. Adaptive Testing and Customized Evaluation

Adaptive testing platforms like ALEKS and Edmentum exemplify how AI personalizes summative assessments. [17] demonstrate that adaptive tests dynamically adjust question difficulty based on real-time student responses, yielding more accurate measures of competence. [18] further explain that these systems pinpoint knowledge gaps, allowing educators to align teaching strategies with actual learning needs. By constructing unique trajectories for each learner, adaptive assessments ensure that evaluations reflect individual capabilities rather than standardized benchmarks.

2.2.3. AI in Examination Supervision and Integrity Assurance

Maintaining academic integrity in online and hybrid settings has driven the adoption of AI-based proctoring solutions. [19] describe platforms like Proctorio and Examity, which use algorithms to monitor eye movements, keystrokes, and ambient sound for suspicious behavior. [20] adds that such surveillance tools are vital in mitigating risks of academic misconduct. While controversial in terms of privacy, these solutions have become increasingly significant in ensuring the credibility of online summative evaluations.

2.2.4. AI in Essay Evaluation and Feedback Mechanisms

AI is also reshaping essay evaluation through systems like WriteLab and Quillionz. [21] demonstrate that WriteLab applies NLP to deliver feedback on grammar and argumentation structure, significantly accelerating the revision process. Similarly, [22] highlights Quillionz, which generates exam questions from instructional texts, helping instructors create comprehensive assessments with less preparation time. Both tools contribute to a more efficient and feedback-rich evaluation cycle, reducing the lag between submission and improvement.

2.2.5. Holistic Learning Analytics Platforms

Platforms such as Brightspace by D2L and SAS EVAAS exemplify holistic use of AI in summative evaluation. [23] notes that these platforms aggregate student data to identify patterns of engagement, enabling teachers to detect at-risk learners early. [24] further show that predictive analytics embedded in such systems allow

institutions to design growth-oriented assessments that promote equity and long-shows academic success.

2.2.6. Plagiarism Identification and Content Authenticity

Finally, AI plays a critical role in verifying originality through plagiarism detection. [25] discusses tools such as Urkund and Unicheck, which scan student submissions against vast databases and the internet to detect overlaps. These technologies are essential in maintaining fairness in written assessments, ensuring academic integrity, and discouraging misconduct.

2.3. Best Practices for AI Integration in Assessment

AI promises transformative benefits in assessment, but its success depends on adopting practices that uphold transparency, fairness, and ethics. [18] stress that both formative and summative uses of AI require frameworks that mitigate risks while enhancing opportunities.

2.3.1. Transparent and Explainable AI

[26] argues that transparency and explainability are critical. Teachers and students must be able to understand how AI tools generate results to maintain trust. Transparent AI systems allow scrutiny of decision-making processes and foster accountability in both formative and summative assessments.

2.3.2. Bias Mitigation and Equity

Bias in AI datasets can perpetuate inequality if not carefully managed. [18] emphasize routine audits of training datasets to ensure fair representation. [27] add that when AI models are built on diverse data, they yield more equitable outcomes. Regular monitoring ensures impartiality in both formative feedback and high-stakes summative results.

2.3.3. Human Oversight and Professional Development

While AI can process vast amounts of data, human expertise remains indispensable. [23] argue that teachers must be trained to interpret AI feedback effectively, using it to refine instruction. [18] further note that human oversight ensures AI complements, rather than replaces, pedagogical judgment.

2.3.4. Balanced Assessment Design

Effective assessment requires balance between AI-driven and traditional approaches. [18] suggest that while AI provides immediate, adaptive feedback, traditional methods such as projects and portfolios capture deeper learning outcomes. A blended approach ensures both reliability and comprehensiveness.

2.4. Advantages of AI in Classroom Assessment

The integration of AI offers numerous advantages for both learners and educators. Adaptive platforms such as ALEKS and DreamBox deliver personalized learning trajectories tailored to individual pace and ability [10]. [28] report that such per-

sonalization reduces student anxiety, increases engagement, and improves mastery.

Immediate feedback is another critical advantage. [29] stresses that tools like Grammarly, WriteLab, and Socrative empower students to correct mistakes instantly, fostering active participation and stronger writing and problem-solving skills. [13] further note that real-time analytics help teachers intervene proactively, improving overall instructional effectiveness.

Educators also benefit as AI reduces administrative burdens. Automation through tools such as Gradescope ensures large-scale grading with high accuracy and consistency, freeing teachers to focus on higher-order tasks such as instructional design and student mentoring. Predictive analytics provide insights into at-risk students, allowing for timely intervention [13].

2.5. Challenges in AI Integration for Equitable and Transparent Assessment

Despite these benefits, challenges remain. [30] emphasize data privacy and security risks, as AI systems often collect sensitive student data vulnerable to breaches. Algorithmic bias is another persistent concern, with underrepresentation in datasets potentially leading to inequitable results [27]. [26] stresses that the opacity of many AI systems undermines trust, particularly in high-stakes contexts. [23] highlights that inadequate infrastructure and limited teacher training often impede successful integration, especially in underfunded schools.

To ensure ethical and effective use, continuous auditing, strong governance, and professional development are essential. Without these safeguards, AI may risk reinforcing inequities rather than reducing them.

3. Methodology

This paper used a rigorous literature review method to analyze and review existing studies on the effective use of AI in the classroom for effective classroom assessments (See **Table 1**). The study laid emphasis on the area of personalization, efficiency and ethical use as the foundational base of the study. The strategy for search for the study was to look for studies that talked about how AI can be used in the classroom for effective formative and summative assessments while laying emphasis on the ethical approach. The study used the following databases while looking for the right studies that aligned with the paper. These were Google scholar, Eric, Jstor and web of science.

3.1. Inclusion

When an article satisfied the following requirements, it was considered for inclusion:

Published between 2013 and 2024; based on peer-reviewed journal articles or conference papers; centered on artificial intelligence tools utilized in elementary, secondary, or higher education; directly relevant to formative or summative assessment practices; addressed one or more of the following: personalization, feedback efficiency, ethics, or teacher implementation.

Table 1. Summary of studies for AI in classroom assessments.

Author(s), Year	Context	Sample	AI Tool(s)	Assessment Type	Key Findings/Outcomes
Akgun & Greenhow (2022)	Higher Education, U.S.	University students	Adaptive platforms, real-time feedback	Formative & Summative	AI improved efficiency and engagement; need for teacher training.
Balfour (2013)	MOOCs	Thousands of online learners	Automated essay scoring	Summative	Scalable essay scoring, but validity and fairness concerns.
Benitez, Gordon, & Olson (2017)	K-12 & HE, conceptual	N/A	Various AI models	Both	Highlighted risks of bias and inequity; called for ethics frameworks.
Chaudhry & Kazim (2022)	Conceptual	N/A	General AI frameworks	Both	Defined AI-human cognition links; emphasized adaptive potential.
Chen, Fan, & He (2020)	K-12, U.S.	Middle school students	ALEKS (adaptive math system)	Formative	Improved mastery and engagement; reduced anxiety via personalization.
Dawson (2016)	Higher Education	Remote exam takers	AI-based proctoring	Summative	Identified cheating strategies; highlighted security gaps in AI proctoring.
Delgado <i>et al.</i> (2020)	Meta-analysis, K-12	Multiple studies	Adaptive learning platforms	Both	AI boosted learning outcomes; effects strongest in well-resourced schools.
Dikli (2019)	Conceptual	N/A	Automated essay scoring	Summative	Reviewed applications; noted potential and limitations for fairness.
Holmes, Bialik, & Fadel (2019)	Global education	N/A	General AI in education	Both	Discussed potential for AI to reshape assessment practices.
Holstein, McLaren, & Aleven (2018)	K-12 classroom	Teachers + students	Orchestration tools	Formative	Real-time AI feedback improved engagement and reduced teacher workload.
Hooda <i>et al.</i> (2022)	Higher Education	College students	AI feedback systems	Formative	Immediate feedback improved performance; scalability shown.
Hwang & Tu (2021)	Online writing courses	Large-scale learners	NLP tools (Grammarly, WriteLab)	Formative	Enhanced writing quality and reduced grading burden.
Ifenthaler & Schumacher (2016)	Higher Education	Review of studies	Learning analytics platforms	Both	Identified need for teacher training and balanced human-AI oversight.
Johnson & Lester (2016)	K-12 simulations	Case-based	ALEKS, Edmentum	Summative	Adaptive testing yielded precise measurement of student ability.
Jomaa (2025)	Higher Education	Writing students	Grammarly, WriteLab	Formative	AI feedback improved writing and translation skills.

Continued

Koedinger <i>et al.</i> (2013)	K-12 and HE	Large-scale data	Predictive analytics	Summative	Highlighted early detection of at-risk students.
McMurtrie (2018)	Higher Education	N/A	General AI in classrooms	Both	Journalistic account; highlighted challenges with teacher adoption.
Murphy (2019)	K-12	Teachers & students	NLP tools (Grammarly)	Formative	Raised privacy and bias concerns; emphasized equity.
Popenici & Kerr (2017)	Higher Education	Conceptual	General AI in teaching	Both	Warned of risks of automation without ethical safeguards.
Selwyn (2019)	Book, global scope	N/A	Conceptual AI frameworks	Both	Raised critical questions on replacing teachers; ethics central.

3.2. Exclusion

The following types of materials are not considered: opinion pieces, editorials, or blog posts that have not been peer-reviewed; articles that are only focused on artificial intelligence in administrative or non-assessment settings; publications written in languages other than English.

3.3. Selection and Analysis

Twenty papers were chosen for the final synthesis after the inclusion criteria were applied, duplicates were removed, and full-text reviews were carried out. A theme analysis was performed on the papers that were chosen.

Within the scope of this review, qualitative theme synthesis was the primary emphasis, with the identification of convergent results and ongoing issues. Two independent reviewers (the first and second authors) conducted the initial coding. Each study was read in full, and segments relevant to AI in classroom assessment were highlighted and coded inductively. Codes captured concepts such as personalization, efficiency, ethical issues, and implementation challenges. Through iterative comparison, these codes were grouped into higher-order categories, which were then refined into four major themes: 1) personalization, efficiency, and scalability; 2) equity and inclusion challenges; 3) ethical and transparency concerns; and 4) methodological gaps in literature.

Limitations

While this systematic review provides valuable insights into the integration of AI in classroom assessment, several limitations must be acknowledged. First, the review included only 20 studies, which narrows the evidence base and may limit generalizability. Second, the review was restricted to English-language publications, which may have excluded relevant studies published in other languages and introduced a potential language bias. Third, as with most systematic reviews, there is the possibility of publication bias, since peer-reviewed journal articles were prioritized over gray literature. Finally, most of the included studies originated from

high-resource educational contexts, leaving gaps in understanding AI's role in low-resource or marginalized settings. These limitations suggest the need for broader, multilingual, and cross-context research to strengthen the evidence base and ensure equity in AI adoption for classroom assessment.

3.4. Prisma Framework

Identification

Records identified through database searching (n = 1240)
 Databases: Google Scholar, ERIC, JSTOR, Web of Science
 Additional records identified through other sources (n = 45)
 Total records identified (n = 1285)
 Duplicates removed (n = 235)
 Records after duplicates removed (n = 1050)

Screening

Records screened (title & abstract) (n = 1050)
 Records excluded (n = 930)

Eligibility

Full-text articles assessed for eligibility (n = 120)
 Full-text articles excluded (n = 100)

Reasons:

- Not peer-reviewed (30)
- Non-assessment focus (25)
- Editorial/opinion (20)
- Other (25)

Included

Studies included in final synthesis (n = 20)

- Conceptual/Review (n = 8)
- Empirical (n = 12)

4. Findings and Discussion

4.1. Theme 1: Personalization, Efficiency, and Scalability

The studies examined indicate that AI-driven assessment systems effectively create tailored learning routes, automate grading, and provide scalable formative feedback to large groups. In high-resource environments, adaptive testing platforms and AI-integrated learning management systems (LMS) adjust difficulty levels in real time, offering tailored feedback that accommodates varied student profiles [31] [32]. Evidence suggests that adaptive algorithms enhance student engagement and reduce achievement disparities, though this depends heavily on adequate infrastructure and teacher training [31] [32]. By contrast, in low-resource environments, these advantages remain largely aspirational. Limited internet connectivity, inadequate device access, and the absence of localized content constrain the effective use of AI's adaptive capabilities [33]. This disparity is further exacerbated by methodological bias in the literature: most empirical data originates from

technologically advanced settings, leaving AI's scalability in disadvantaged educational contexts underexplored.

4.2. Theme 2: Equity and Inclusion Challenges

AI is often portrayed as a democratizing force in education; however, the evidence reviewed suggests that its implementation can perpetuate, rather than eliminate, existing inequalities. In affluent systems, AI-based assessments align with curriculum standards, supported by sustained teacher professional development and favorable policy frameworks [34]. In contrast, in low-resource contexts, limited teacher training in AI literacy and weak digital infrastructure lead to surface-level adoption and underutilization of AI capabilities. Studies warn that without intentional equity frameworks, AI integration risks deepening digital divides [12]. Furthermore, there is limited involvement of teacher and student voices from disadvantaged contexts, which reduces the cultural validity of current evaluations. This gap underscores the urgent need for participatory, context-specific approaches to guide AI adoption equitably.

4.3. Theme 3: Ethical and Transparency Concerns

Ethical issues such as algorithmic bias, opaque decision-making, and risks to data privacy were recurring concerns across studies. In well-resourced settings, regulatory frameworks and institutional safeguards require greater algorithmic transparency and stakeholder consultation [35]. In contrast, such protections are largely absent in low-resource environments, where weak regulatory oversight leaves schools more vulnerable. Studies highlight the lack of standardized explainability tools, insufficient culturally relevant metrics, and restricted access to training datasets. These shortcomings are particularly harmful in linguistically and culturally diverse schools, where algorithmic decisions may fail to reflect community norms or values. Without clear review and accountability processes, trust in AI-driven assessment remains fragile, especially in underregulated settings.

4.4. Theme 4: Methodological Gaps and Research Imbalances

Another notable limitation of the current evidence base is its reliance on pilot studies and small-scale case analyses in well-funded systems. Longitudinal investigations of sustained impacts are rare, while cross-cultural or cross-socioeconomic comparisons remain scarce. Few studies employ robust mixed-methods designs that triangulate quantitative performance data with qualitative insights from learners and educators [36]. Many evaluations are conducted under controlled conditions, which fail to capture the complexities of everyday classroom realities in disadvantaged contexts. This methodological narrowness undermines external validity and restricts understanding of AI's adaptability to diverse cultural, linguistic, and infrastructural conditions.

Taken together, the findings suggest that while AI has the potential to improve personalization, efficiency, and scalability in classroom assessment, its benefits are

distributed unevenly. In high-resource environments, strong infrastructure, teacher expertise, and policy alignment foster successful implementation. In contrast, low-resource settings face structural barriers that may negate potential advantages and exacerbate inequities. Where regulatory oversight is weak, ethical concerns related to bias, transparency, and data security become especially acute. The methodological gaps in existing research further complicate efforts to generalize findings or design context-sensitive implementation strategies. Ultimately, realizing AI's potential in education requires equity-centered approaches that prioritize infrastructure investment, culturally responsive design, teacher capacity-building, and strong ethical governance. Absent these measures, AI risks becoming a mechanism that entrenches existing disparities rather than a transformative tool for educational equity.

5. Conclusions

This research demonstrates that Artificial Intelligence is transforming classroom assessment by improving customization, efficiency, and scalability in both formative and summative settings. Evidence suggests that AI solutions, like adaptive learning platforms and automated scoring systems, provide quantifiable advantages in student engagement, precise feedback, and grading uniformity. The benefits are particularly evident in high-resource settings, where strong infrastructure, professional development, and policy support facilitate ongoing integration. Conversely, low-resource environments have substantial obstacles, such as restricted access to devices and connections, inadequate teacher training, and fragile data governance structures. In the absence of intentional measures, these discrepancies threaten to solidify rather than diminish educational inequities.

Ethical and technological challenges—such as algorithmic bias, insufficient transparency, and data privacy issues, are fundamental factors that determine whether AI functions as a mechanism for equality or exclusion. The literature examined indicates a dual perspective: optimism over AI's capacity to revolutionize assessment into a more learner-centric and adaptable methodology, and caution in acknowledging that unthinking implementation may perpetuate systemic biases and erode trust.

This synthesis is opportune. As global education systems contend with post-pandemic recovery and heightened demands for tailored, competency-based learning, AI in assessment has transitioned from a theoretical concept to a practical implementation. The issue is not the use of AI, but rather the methods of implementation that ensure justice, transparency, and inclusion. Addressing the disparate data base, especially the lack of research from low-resource settings, and investing in longitudinal, comparative studies will be essential to ensure that AI-driven evaluations realize their potential without undermining educational fairness.

This research emphasizes that the effective incorporation of AI into classroom assessment requires not just technologies but also governance, capacity develop-

ment, and a dedication to fairness throughout the adoption process, by contextualizing the results within technical and socio-political realities.

This study illustrates that Artificial Intelligence is transforming classroom assessment by facilitating enhanced customization, efficiency, and scalability in both formative and summative settings. High-resource settings enjoy advanced infrastructure, comprehensive teacher training, and supportive policy frameworks, whereas low-resource environments encounter obstacles like limited connectivity, device shortages, and inadequate professional development, potentially exacerbating educational inequalities. Ethical challenges, specifically algorithmic prejudice, data privacy issues, and insufficient transparency—are essential in assessing whether AI adoption fosters inclusiveness or reinforces structural injustices. The evidence base is inconsistent, with a significant lack of research from marginalized settings, highlighting the need for equity-centered implementation. To rectify these deficiencies, policymakers and educators must implement equity-focused AI frameworks that evaluate bias and cultural pertinence prior to deployment, invest in professional development that transcends technical competencies to encompass critical analysis of AI outputs, enforce transparency and explainability standards for AI tools to cultivate trust, and enhance infrastructure in resource-limited environments through cross-sector collaborations. As educational systems address post-pandemic recovery and transition to competency-based learning, this synthesis underscores that AI's function in assessment is now practical rather than theoretical, necessitating intentional, equitable, and transparent implementation by educators and policymakers globally.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] Chaudhry, M.A. and Kazim, S. (2022) Artificial Intelligence in Computer Science: Bridging Human and Machine Cognition. *International Journal of Computer Science and AI*, **10**, 23-31.
- [2] Holmes, W., Bialik, M. and Fadel, C. (2019) Artificial Intelligence in Education: Promises and Implications for Teaching and Learning. Center for Curriculum Redesign.
- [3] Krueger, K. (2017) AI's Promise in K-12 Education: Supporting Students and Educators. *Education Week*, **36**, 9-10.
- [4] McMurtrie, B. (2018) How AI Could Be the Next Big Thing in Education. *Chronicle of Higher Education*, **64**, 14-16.
- [5] Akgun, B. and Greenhow, C. (2022) The Integration of AI in Education: Classroom Assessment and Student Engagement. *Journal of Educational Technology*, **33**, 245-258.
- [6] Delgado, A., Wardlow, L., Olin, S. and McKnight, K. (2020) The Effects of Using Adaptive Learning Technology on Student Learning Outcomes: A Meta-Analysis. *Educational Technology Research and Development*, **68**, 2161-2182.
<https://doi.org/10.1007/s11423-020-09807-4>

- [7] Murphy, A. (2019) AI in Education: A New Frontier in Formative Assessment and Feedback. *Educational Leadership*, **76**, 45-49.
- [8] Popenici, S.A.D. and Kerr, S. (2017) Exploring the Impact of Artificial Intelligence on Teaching and Learning in Higher Education. *Research and Practice in Technology Enhanced Learning*, **12**, 22-37. <https://doi.org/10.1186/s41039-017-0062-8>
- [9] Chen, S., Fan, W. and He, L. (2020) Adaptive Learning Systems for Personalized Education: Insights from the Case of ALEKS. *Educational Technology Research and Development*, **68**, 1-18. <https://doi.org/10.1007/s11423-019-09713-7>
- [10] Hwang, G.J. and Tu, Y.F. (2021) Roles and Applications of Artificial Intelligence in Education. *Computers & Education: Artificial Intelligence*, **2**, Article ID: 100015.
- [11] Zawacki-Richter, O., Marín, V.I., Bond, M. and Gouverneur, F. (2019) Systematic Review of Research on Artificial Intelligence Applications in Higher Education—Where Are the Educators? *International Journal of Educational Technology in Higher Education*, **16**, 1-27. <https://doi.org/10.1186/s41239-019-0171-0>
- [12] Holstein, K., McLaren, B.M. and Alevan, V. (2018) Student Learning Benefits of a Real-Time Classroom Orchestration Tool Supported by Teacher AI. *Proceedings of the International Conference on Learning Analytics & Knowledge (LAK)*, Vol. 10, 11-15. <https://doi.org/10.1145/3170358.3170361>
- [13] Balfour, S.P. (2013) Assessing Writing in MOOCs: Automated Essay Scoring and Calibrated Peer Review. *Research & Practice in Assessment*, **8**, 40-48.
- [14] Strobl, C., et al. (2019) Machine Learning for Automated Essay Scoring: A Review. *Journal of Educational Computing Research*, **57**, 331-353. <https://doi.org/10.1177/0735633117741520>
- [15] Johnson, W.L. and Lester, J.C. (2015) Face-to-Face Interaction with Pedagogical Agents, Twenty Years Later. *International Journal of Artificial Intelligence in Education*, **26**, 25-36. <https://doi.org/10.1007/s40593-015-0065-9>
- [16] Shute, V.J. and Rahimi, S. (2017) Review of Computer-Based Assessment for Learning in Elementary and Secondary Education. *Journal of Computer Assisted Learning*, **33**, 1-19. <https://doi.org/10.1111/jcal.12172>
- [17] Stone, D., et al. (2020) The Role of AI in Remote Proctoring and Exam Security. *Educational Research Review*, **31**, Article ID: 100334.
- [18] Dawson, P. (2015) Five Ways to Hack and Cheat with Bring-Your-Own-Device Electronic Examinations. *British Journal of Educational Technology*, **47**, 592-600. <https://doi.org/10.1111/bjet.12246>
- [19] Li, H., Yang, S. and Sun, Y. (2021) Evaluating Automated Feedback Mechanisms in Writing Assessments. *Journal of Educational Measurement*, **58**, 222-240.
- [20] Dikli, S. (2019) Automated Essay Scoring: Applications to Educational Technology. *The Journal of Technology, Learning, and Assessment*, **3**, 1-20.
- [21] Ifenthaler, D. and Schumacher, C. (2016) Learning Analytics in Higher Education: A Literature Review. *British Journal of Educational Technology*, **47**, 924-937.
- [22] Koedinger, K.R., et al. (2013) Predictive Analytics in Education: Methods and Applications. *International Journal of Learning Analytics and Artificial Intelligence in Education*, **7**, 3-19.
- [23] Dawson, S. (2016) Five Ways to Hack and Cheat with Electronic Examinations. *British Journal of Educational Technology*, **47**, 592-600.
- [24] Walker, J. (2010) Student Plagiarism in Universities: Learning for Better Assessment in Higher Education. *Studies in Higher Education*, **35**, 387-406. <https://doi.org/10.1080/03075070903062892>

- [25] Sharma, D., Jones, R. and Smith, K. (2020) The Ethics of AI in Education: Privacy and Security Concerns. *Journal of Educational Technology & Society*, **23**, 75-85.
- [26] Benitez, K., Gordon, A. and Olson, J. (2017) AI in Education: The Importance of Ethics and Bias Management. *AIED Journal*, **25**, 112-128.
- [27] Selwyn, N. (2019) *Should Robots Replace Teachers? AI and the Future of Education*. John Wiley & Sons.
- [28] Jomaa, N.J. (2025) Corrective Feedback of AI Tools in Writing and Translation. In: *Using AI Tools in Text Analysis, Simplification, Classification, and Synthesis*, IGI Global, 293-318. <https://doi.org/10.4018/979-8-3693-9511-0.ch010>
- [29] Hooda, M., Rana, C., Dahiya, O., Rizwan, A. and Hossain, M.S. (2022) Artificial Intelligence for Assessment and Feedback to Enhance Student Success in Higher Education. *Mathematical Problems in Engineering*, **2022**, Article ID: 5215722. <https://doi.org/10.1155/2022/5215722>
- [30] West, R., Hoopes, R. and Berrett, D. (2020) The Future of AI in Summative Assessments: Implications and Innovations. *Educational Assessment*, **25**, 253-272. <https://doi.org/10.1080/10627197.2020.1832680>
- [31] Holmes, W. and Tuomi, I. (2022) State of the Art and Practice in AI in Education. *European Journal of Education*, **57**, 542-570. <https://doi.org/10.1111/ejed.12533>
- [32] Olanrewaju, A.S. (2024) The Nexus of Human Communication and AI in the Workplace. *Journal of Communication and Management*, **3**, 260-264. <https://doi.org/10.58966/jcm20243311>
- [33] Heffernan, N.T. and Heffernan, C.L. (2014) The Assessments 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**, 470-497. <https://doi.org/10.1007/s40593-014-0024-x>
- [34] Luckin, R., Holmes, W., Griffiths, M. and Forcier, L.B. (2016) *Intelligence Unleashed: An Argument for AI in Education*. Pearson.
- [35] Williamson, B. and Piattoeva, N. (2022) Education Governance and Datafication. *Education and Information Technologies*, **27**, 3515-3531.
- [36] Ifenthaler, D. and Yau, J.Y. (2020) Reflections on Different Learning Analytics Indicators for Supporting Study Success. *International Journal of Learning Analytics and Artificial Intelligence for Education*, **2**, 4-23. <https://doi.org/10.3991/ijai.v2i2.15639>