International Summit on Science Technology and Humanity
ISETH 2025
ISSN: 2615-1588 (online)

The Digital Paradox in Schools: Navigating The Intersection between Artificial Intelligence and Academic Performance

Nurul Fahri Friyanda[1*], Dartim[1]
[1]Department of Islamic Education, Faculty Islamic studies, Universitas Muhammadiyah Surakarta, Surakarta, Indonesia

Abstract
The acceleration of the digitalization of education and the introduction of artificial intelligence (AI) into the classroom have created a "digital paradox": technologies designed to improve the quality of learning can simultaneously have opposite impacts, such as decreased motivation, weakened deep understanding, threats to academic integrity, and inequality in learning outcomes. This study aims to map and synthesize empirical evidence on the relationship between the use of AI/digital education and academic performance, as well as identify the conditions that make its impact stronger or weaker. The method used is a PRISMA-based Systematic Literature Review (SLR) with Scopus as a single database. The search uses structured Boolean strings in the range of 2021–2025, open access journal articles, Social Sciences/Arts and Humanities, and English. From the initial 562 articles, it was filtered down to 38, and through relevance screening based on title–abstract–content verification reading, 9 studies that met the inclusion criteria were obtained. Narrative synthesis shows evidence that the impact of AI on academic performance is contextual and not always consistent: some studies highlight benefits (learning support, personalization, early warning, acceleration of assessment), while others emphasize affective trade-offs (interest/motivation), risks of cognitive "shortcut" behavior, and validity-fairness issues in digital assessments. These findings confirm that the impact of AI is more determined by task–AI fit, AI literacy and ethics, assessment design that assesses thought processes, and equity-sensitive school governance.
Keywords: *artificial intelligence; digital education; digital paradox; academic performance; systematic literature review.*

Introduction: The Digital Paradox in Schools

The acceleration of the digitalization of education in the last decade has radically changed the learning ecosystem in schools. Globally, publications on AI-based adaptive learning jumped from just 1 article in 1990 to 636 in 2023, showing an explosion of interest as well as the adoption of smart technologies in formal education (Dogan et al., 2023; Gligorea et al., 2023; Hamal et al., 2022; Strielkowski et al., 2024). Another bibliometric study of 3518 publications themed "adaptive learning" and "AI" confirms that this surge was strongly triggered by the post-pandemic "digital surge" COVID19 which forced schools and universities to switch to AI-integrated digital learning ecosystems (Strielkowski et al., 2024). In the higher education sector, a systematic review found 64 studies that specifically analyzed the impact of AI adoption on academic performance in open and distance learning, with a general tendency that AI is able to enrich learning experiences while uncovering a non-uniform pattern of impacts on academic achievement (Adewale et al., 2024). At the same time, generative AI like ChatGPT dominated 207 studies in one major review, and was identified as a "game changer" that is now penetrating the classroom through smart tutors, automated assessments, and academic writing support (Bahroun et al., 2023). This fact shows that in schools, AI is no longer just an add-on, but is starting to reshape the way students learn, teachers design assignments, and institutions assess learning outcomes.

Conceptually, artificial intelligence in schools in this study is understood as the use of AI-based applications and features—including generative AI, intelligent tutoring systems, adaptive learning platforms, and automatic correction tools—by students and teachers to help the learning process, task completion, and evaluation (Yarlagadda, 2025). Academic performance refers to measurable indicators of learning outcomes, such as subject scores, exam or assessment scores, completeness of curriculum competencies, and the quality of assignment products (e.g. learning design, reports, or essays)(Mariñoso et al., 2025). Meanwhile, the digital paradox in this context describes a condition when technology that is normatively designed to improve the quality of learning actually produces a double impact: on the one hand it strengthens learning outcomes, on the other hand it has the potential to weaken the depth of understanding and academic integrity, depending on usage patterns, school regulations, and student characteristics (Zhai et al., 2024). The scope of the study is focused on the use of AI by students in the school/university environment and its implications on academic performance, taking into account the relationship with the design of teacher assignments and school policies.

In terms of opportunities, the educational literature shows that AI has a strong pedagogical basis to improve academic performance. Adaptive learning and intelligent tutoring systems are able to analyze student performance data on an ongoing basis to adjust the level of difficulty, speed, and form of material presentation so that it is more aligned with individual needs(Yarlagadda, 2025). With machine learning algorithms, adaptive platforms can provide instant feedback, offer personalized repetitive exercises, and automatically direct students to remedial materials, which have been shown to correlate with improved learning achievement and retention (R et al., 2025). A quasi-experimental study of 238 early childhood education students showed that the group that used AI to design math learning situations achieved significantly better assignment results than the group that did not use AI, marking the real potential of AI in enriching the quality of academic products (Mariñoso et al., 2025). In addition, the bibliometric review generative AI notes the strong role of AI

[*]Corresponding author: g000239147@student.ums.ac.id

in automated assessment, 24/7 self-paced learning support, and large-scale personalization that theoretically supports learning differentiation and the expansion of effective learning time for students (Kamalov et al., 2023).

However, the other side of this spectrum presents no less significant risks. A systematic review of over-reliance on AI dialogue systems found that when students tended to accept AI recommendations uncritically, their analytical thinking, reasoning, and decision-making abilities were eroded; AI becomes a "quick heuristic" that replaces deep thinking processes (Zhai et al., 2024). A qualitative study with college students in Pakistan reported that the use of generative AI to complete assignments led some students to only paraphrase the AI answers and submit them as their own work, resulting in active information processing and decreased creativity (Majeed et al., 2025). Similar concerns are seen in educators who believe algorithm-based writing can encourage laziness, hinder the development of thinking skills, and blur the boundaries of personal contribution in academic writing (Gustilo et al., 2024). In the realm of assessment, a scoping study of generative AI highlights the potential for increased plagiarism, ghostwriting, and erosion of academic integrity if institutions do not revise their assessment designs and policies (Perkins et al., 2023). Furthermore, various reviews emphasized the risk of gaps due to differences in device access, digital literacy/AI, and institutional support, so that the benefits of AI are not felt equally (Aslam et al., 2025).

It is this tension between opportunity and risk that gives rise to the "intersection" of the digital paradox as reflected in the title of the study. At one pole, AI functions as a learning support, a cognitive partner that enriches the learning process—providing scaffolding, examples, feedback, and adaptive training paths that encourage concept mastery and learning independence (Yarlagadda, 2025). At the other pole, AI operates as a task replacement, when students primarily use it to complete tasks instantly: generating text, solutions, or designs that are then only slightly modified before being collected (Khan & Irfan, 2025). Another intersection arises in the dichotomy of using etistetrics—which are governed by learning objectives, rubrics, and AI literacy guides—versus the use of pragmatism that focuses on the speed of completing tasks and obtaining grades (Francis et al., 2025). Recent literature suggests that these differences in how they are used may well explain why in some contexts AI appears to improve academic performance, while in others it weakens it (Mariñoso et al., 2025).

Empirically, the evidence on the impact of AI on academic performance is still mixed and inconsistent. On the one hand, there are studies that show improved grades, task quality, and more accurate prediction of achievement through AI-based learning analytics (Adewale et al., 2024). On the other hand, a cross-regional survey of 300 students and lecturers found a weak and inconsistently positive correlation between the use of AI and learning outcomes and instructional quality, suggesting that the potential of AI is often overestimated and highly dependent on institutional readiness, lecturer competence, and student motivation (Aslam et al., 2025). Numerous large reviews also confirm that although AI applications have been widely mapped (intelligent tutoring, personalization, assessment, analytics), research that systematically links AI usage patterns to measurable academic outcomes—as well as differentiates the type of use (learning concepts vs. completing assignments), intensity, institutional context, and internal student factors—is still limited (Bahroun et al., 2023). Systematic reviews in the context of distance learning, for example, explicitly mention the absence of a process framework capable of predicting more precisely how AI impacts academic performance, including the role of moderation factors such as gender and region (Adewale et al., 2024). Similarly, reviews on AI in higher education call for longitudinal studies and models that are able to explain the conditions under which AI actually results in increased achievement and when it does not.

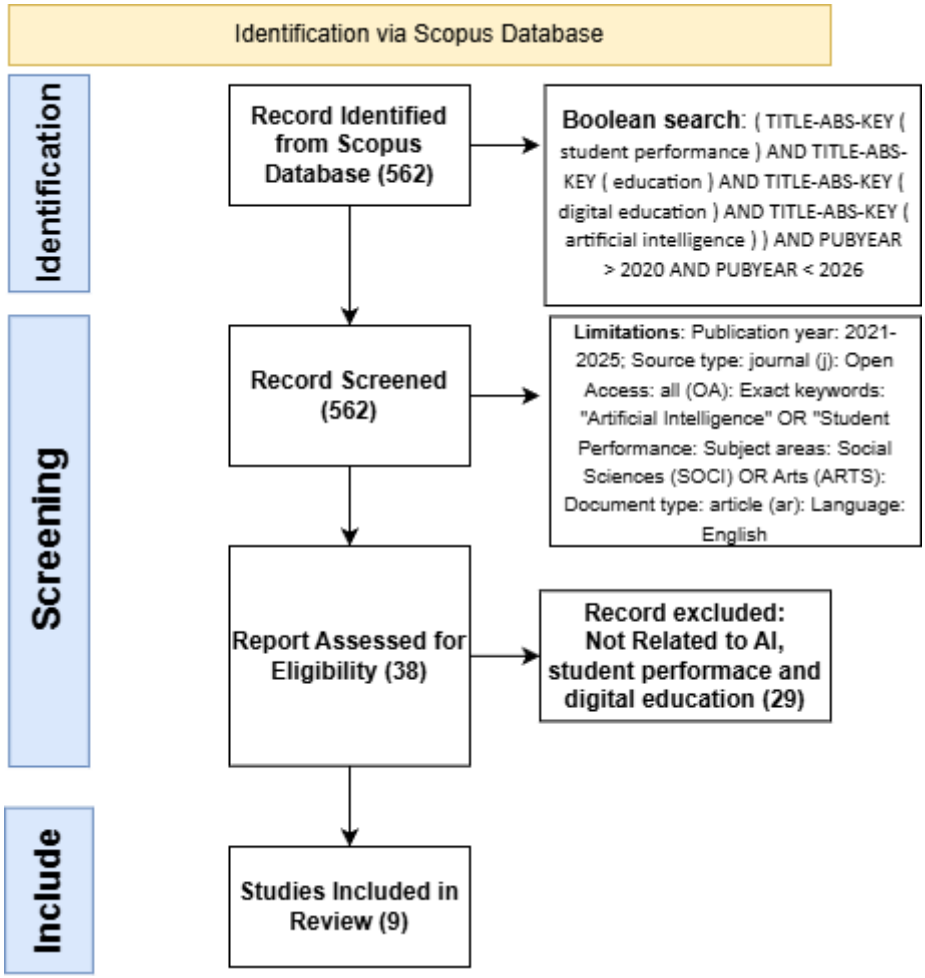
These conditions underscore the need for research that maps and navigates key variables at the intersection between AI use and academic performance in schools. Conceptually, the main relationship in focus is the use of academic performance AI, with particular attention to how the intensity, objectives, and types of use of AI (support vs replacement) relate to score indicators, assessment scores, competency completeness, and task quality (Adewale et al., 2024). In it, a number of factors are estimated to play a mediator, such as independent learning and self-regulated learning strategies, the quality of conceptual understanding, engagement in learning activities, and the cognitive burden experienced by students in interacting with AI. On the other hand, important moderators that need to be taken into account include AI literacy and students' digital ethics, pedagogical support and teacher supervision, school policies and academic culture related to the use of AI, types of assignments (high-level thinking/HOTS vs routine), and access to technology infrastructure (Kamalov et al., 2023). By mapping the configuration of these mediators and moderators, the research on "The Digital Paradox in Schools: Navigating the Intersection between Artificial Intelligence and Academic Performance" is expected to be able to explain more sharply when and under what conditions AI functions as a lever for academic performance, and when it is a source of declining learning quality and academic integrity.

Methodology

This study uses the Systematic Literature Review (SLR) approach to collect, select, and synthesize scientific evidence regarding the relationship between artificial intelligence (AI), digital education, and student performance in the field of Social Sciences and Arts and Humanities studies. All stages of study selection and reporting are carried out transparently by referring to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework, so that the process of identification, screening, eligibility determination, and inclusion of articles can be systematically traced back through a clear flow.

The data source used in this study is Scopus as a single database. The search strategy is designed using the following Boolean strings: (TITLE-ABS-KEY (student performance) AND TITLE-ABS-KEY (education) AND TITLE-ABS-KEY (digital education) AND TITLE-ABS-KEY (artificial intelligence)) AND PUBYEAR > 2020 AND PUBYEAR < 2026 AND (LIMIT-TO (SRCTYPE , "j")) AND (LIMIT-TO (OA , "all")) AND (LIMIT-TO (EXACTKEYWORD , "Artificial Intelligence") OR LIMIT-TO (EXACTKEYWORD , "Student Performance")) AND (LIMIT-TO (SUBJAREA , "SOCI") OR LIMIT-TO (SUBJAREA , "ARTS")) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (LANGUAGE , "English"))

Diagram 1. Reporting Items for Systematic Reviews and Meta-Analyses Flow



The inclusion criteria are strictly set so that only articles that are relevant and in line with the research objectives are maintained. The article must be an empirical study, meaning that the study presents the use of data (primary or secondary) and an explicit method of analysis to evaluate the phenomenon being studied. In addition, the article must be open access, indexed by Scopus in the field of Social Sciences or Arts and Humanities, in the form of a journal article, and substantially discuss the relationship or influence of AI and/or digital education on student performance through measurable learning performance indicators (e.g. grades, test scores, competency achievements, or similar academic outcomes). Articles are issued if they are in the form of a purely conceptual study without data testing, do not include the category of documents that are determined, are outside the subject area of SOCI/ARTS, or do not place AI/digital education and student performance as the focus of relevant analytical relationships.

The article selection process follows the PRISMA stage. At the identification stage, the initial search results on Scopus yielded 562 articles. Furthermore, the articles were filtered through the application of inclusion criteria attached to the search filter (e.g. open access, subject area, document type, source type, language, year range, and keyword accuracy), reducing the number of articles to 38. The next stage is follow-up screening to ensure a tighter fit with the focus of the discussion, which is to verify that the article is truly an empirical study and directly discuss AI/digital education in relation to student performance according to the scope of the title and the purpose of the research. Through relevance screening based on reading the title, abstract, and verifying the content of the article, the number of studies that met all the criteria and was judged to be most appropriate for the discussion ended up being 9 articles.

Once the included study is established, structured data extraction is carried out to ensure consistent synthesis. The information collected included the identity of the study (author, year, title, and context), research design and method, sample characteristics and educational setting, how to operationalize AI and digital education, student performance indicators used, analysis techniques applied, and key findings related to the direction and pattern of intervariable relationships. The results of the extraction are then synthesized narratively (narrative synthesis) considering the possibility of variations in the research design, context, and form of learning performance measurement in each article. With this procedure, the methodology of this study ensures that the analyzed articles not only meet the administrative requirements (open access and subject area), but also meet the substantive requirements as relevant empirical evidence to answer the focus of the study.

Results and Discussion

Trends and gaps in research results related to the digital paradox in schools

The digital paradox in schools arises when AI, which is designed to improve learning, actually gives rise to negative effects such as decreased motivation and inequality of outcomes. Three main articles illustrate this paradox through: (1) attitudes towards AI, (2) the impact of AI adaptive feedback on the learning process, and (3) the impact of digitalization on assessment and fairness.

Attitudes towards AI are often ambivalent: on the one hand it is considered to help the task, but it also raises concerns. Therefore, digital competence becomes important, with an emphasis on critical thinking and conscious use (Tsenov & Bakracheva, 2025). This ambivalence indicates that although AI is considered necessary, its use remains in doubt.

In adaptive feedback, Bauer et al. found that while AI is expected to improve performance, the results are not significant; interest in adaptive feedback is declining, because it directs too much or highlights errors, interfering with autonomy and competence (Bauer et al., 2025). This suggests that technology evaluations should consider affective impacts, not just cognitive outcomes.

In the realm of assessment, Yildirim-Erbasli & Bulut point out that in low-stakes digital assessments, students tend to answer origins, undermining the validity of scores. They propose conversation-based assessment (CBA) to increase engagement, although there are still fairness issues, as students with certain technology skills may benefit more (Yildirim-erbasli & Bulut, 2023).

Research trends show a shift from "technology as a solution" to "socio-technical systems" influenced by attitudes and literacy (Tsenov & Bakracheva, 2025), from single indicators (values) to multidimensional evaluations including motivation (Bauer et al., 2025), as well as from efficiency to quality participation and fairness in assessments (Yildirim-erbasli & Bulut, 2023).

Existing research gaps include a lack of a cross-actor framework to understand attitudes towards AI as a whole (Tsenov & Bakracheva, 2025) and a lack of empirical evidence in the context of primary education, which requires contextual and longitudinal studies (Bauer et al., 2025).

The paradox is also seen in the pedagogical design and technology used. Aslan et al. show how concerns about screen time can be addressed with a blended physical–digital design that increases engagement and physical activity, but long-term research is still needed (Aslan et al., 2024). Ilić et al. map out a shift towards intelligent techniques in e-learning, even though the implementation of personalization requires uneven data and institutional capacity, causing a paradox (Ilić et al., 2023). Jokhan et al. show how AI is used for performance prediction, but predictions can be biased if digital access is limited (Jokhan et al., 2022). Leite tested generative chatbots in chemical learning, emphasizing the need for critical use and literacy in interacting with models (Leite, 2024).

In classroom practice and assessment, Mwakalinga & Mabilika show that although students see the benefits of AI, educators are concerned about declining comprehension and critical thinking, requiring ethical guidelines and digital literacy (Mwakalinga & Mabilika, 2025). Terrazas-Arellanes et al. show that AI in science assessments can speed up assessments and reduce bias, but its integration requires adequate teacher training and governance (Terrazas-arellanes et al., 2025).

In synthesis, the digital paradox in schools is becoming clearer when efficiency and scalability come at the expense of the quality of the learning experience and fairness. There are gaps in long-term evidence and non-academic evaluations, which need to be addressed to understand this paradox as a matter of design, governance, and justice, not just a pro-cons of technology (Aslan et al., 2024; Ilić et al., 2023; Jokhan et al., 2022; Leite, 2024; Mwakalinga & Mabilika, 2025; Terrazas-arellanes et al., 2025).

Table 1. collection of inclusion journals

Authors	Title	Year	Journal		Author Affiliation	Subject or Educational Level
Seyma N. Yildirim-Erbasli; Okan Bulut	*Conversation-based assessment: A novel approach to boosting test-taking effort in digital formative assessment*	2023	*Computers and Education: Artificial Intelligence*	Canada	(a) Department of Psychology, Concordia University of Edmonton, Edmonton, AB, Canada; (b) Centre for Research in Applied Measurement and Evaluation, University of Alberta, Edmonton, AB, Canada	formative assessment, digital assessment, conversational agents/AI; for **K–12 and higher education** contexts (general, not subject specific)
M.Y. Tsenov; M.A.S. Bakracheva	*Attitudes towards artificial intelligence in professional and personal life*	2025	*The Education and Science Journal*	Bulgaria	Sofia University "St. Kliment Ohridski", Sofia, Bulgaria	Sample: **125** respondents (students + professionals); there are **FG/interviews** including school & university teachers (dominant context: higher education + professional)
Elisabeth Bauer; Constanze Richters; Amadeus J. Pickal; Moritz Klippert; Michael Sailer; Matthias Stadler	*Effects of AI-generated adaptive feedback on statistical skills and interest in statistics: A field experiment in higher education*	2025	*British Journal of Educational Technology*	Germany	University of Augsburg (Learning Analytics & EDM), Augsburg, Germany; LMU Munich (Institute of Medical Education), Munich, Germany	**Higher education**: S1 students in the first semester of the Educational Sciences program (field experiment in statistics/research methods)
Bruno Silva Leite	*Generative Artificial Intelligence in chemistry teaching: ChatGPT, Gemini, and Copilot's content responses*	2024	*Journal of Applied Learning & Teaching* (Vol.7 No.2)	Brazil	Department of Education, Federal Rural University of Pernambuco, Pernambuco, Brazil	**Chemistry teaching** (analysis of the accuracy of GenAI's response to basic chemistry concepts; focus on content/concepts, not specific classroom tests).
Sinem Aslan; Lenitra M. Durham; Nese Alyuz; Eda Okur; Sangita Sharma; Celal Savur; Lama Nachman	*Immersive multi-modal pedagogical conversational artificial intelligence for early childhood education: An exploratory case study in the wild*	2024	*Computers and Education: Artificial Intelligence* (Vol.6, 100220)	USA	Intel Labs, USA	**Early childhood/elementary (Grade 1)** – primary school context, math material **ones & tens** (number and operations in base ten).

Angela Jokhan; Aneesh A. Chand; Vineet Singh; Kabir A. Mamun	*Increased Digital Resource Consumption in Higher Educational Institutions and the Artificial Intelligence Role in Informing Decisions Related to Student Performance*	2022	*Sustainability*	Fiji	Ministry of Education, Heritage and Arts (Suva, Fiji); School of Information Technology, Engineering, Mathematics and Physics—STEMP (Suva, Fiji); University of the South Pacific/USP (Suva, Fiji)	**Higher education**: predicting student performance in **the first-year IT literacy course** at the University of the South Pacific (USP)
Miloš Ilić; Vladimir Mikić; Lazar Kopanja; Boban Vesin	*Intelligent techniques in e-learning: a literature review*	2023	*Artificial Intelligence Review*	Serbia; Norway	Alfa BK University (Belgrade, Serbia); University of South-Eastern Norway (Vestfold, Norway); Norwegian University of Science and Technology/NTNU (Trondheim, Norway)	**E-learning (cross-level/general);** It does not focus on one domain, but the majority of the systems reviewed come from **computer science education**, as well as languages, mathematics, physics, biology, etc.
Shadrack Ernest Mwakalinga; Filbert Athanas Mabilika	*Perceptions, pitfalls, and proposals for the ethical use of artificial intelligence in the classroom: a case study of students and educators (teachers' and lecturers')*	2025	*Cogent Education*	Tanzania	Catholic University of Mbeya, Mbeya, Tanzania	**Secondary schools & universities**; Stakeholders: students, teachers (secondary), lecturers (university)
Fatima Elvira Terrazas-Arellanes; Lisa Strycker; Giani Gabriel Alvez; Bailey Miller; Kathryn Vargas	*Promoting Agency Among Upper Elementary School Teachers and Students with an Artificial Intelligence Machine Learning System to Score Performance-Based Science Assessments*	2025	*Education Sciences*	USA	College of Education, University of Oregon, Eugene, OR, USA	**Upper elementary (Grades 3–5);** science performance-based assessment (constructed responses)

Based on the table (9 publications), the most visible pattern is the surge in research focus in 2025. Of the total 9 studies, 4 will be published in 2025, while in 2024 there will be 2, in 2023 there will be 2, and in 2022 only 1. This sends a simple but important signal: the topic of AI and digital technology in education is being "upgraded" from discourse to research areas that are increasingly routinely produced, especially when the issue of GenAI (e.g. ChatGPT et al.), adaptive feedback, and AI-based assessments are starting to touch learning practices and policies.

In terms of themes, this data is not just about "using AI to be cool", but moving to a more operational issue: assessment and feedback are the most dominant clusters. There are studies on *conversation-based assessment*, AI-based adaptive feedback to improve statistical and *inference skills*, and the use of AI in performance-based assessment in science. This shows a fairly consistent global trend: AI is used not only as a learning medium, but as a measuring and directing machine for learning (assessment + feedback) that can personalize the learning experience. At the same time, another more "front-class" path has emerged: the utilization of immersive multimodal conversational AI for early learners (e.g. Grade 1 math) and GenAI studies for chemistry learning that assess the accuracy of model responses. That is, research is starting to press on two sides at once: the pedagogical effectiveness and reliability of AI content.

At the same time, research is still focused on higher education: about 4 out of 9 studies focused predominantly on the university/student context (e.g., statistics, IT literacy, attitudes towards AI in students/professionals). Primary education

also emerged (about 2 studies, Grade 1 and Grades 3–5), while the rest were cross-level/general (e.g. literature review e-learning, or ethics studies involving secondary schools and universities). The interpretation is quite clear: campuses often become "laboratories" because of data access, ease of experimentation, and infrastructure readiness. As a result, the evidence at the elementary-secondary school level is still relatively thin, whereas there implementation challenges (teacher readiness, classroom controls, device policies, digital literacy) are usually more brutal.

From the perspective of country distribution, his research contributions are fragmented but global: the United States appears most often (2 publications), then other countries 1 publication each (Canada, Germany, Bulgaria, Brazil, Fiji, Tanzania, as well as the Serbia–Norway collaboration). This illustrates that the issue of educational AI does not belong to one region: there is a strong Europe in learning experiments, North America in conversational/multimodal + assessment, South America in GenAI for chemistry teaching, Oceania in the issue of digital resource consumption in educational institutions, and Africa in the discussion of ethics, perceptions, pitfalls, and proposed use of AI. In other words, the research trends are both towards AI, but the "anxiety" of each region is different: some pursue improving learning performance, some take care of the costs/digital footprint of institutions, and some hold back the pace of technology with ethical brakes.

The interpretive conclusion: this dataset shows a transformation from "digitalization of learning" to more specific AI-ization: *intelligent assessment*, *adaptive feedback*, *conversational* tutoring, and evaluation of GenAI's reliability in subject matter. However, there are perceived gaps: (1) the evidence for implementation in secondary education is still less dense than in higher education, (2) ethical themes have emerged but are still more in the form of perceptions and proposals than tested implementation frameworks, (3) the focus on effectiveness often measures learning outcomes, but does not always close the more important questions: how governance, validity of assessments, and long-term impacts on learning independence. So, the direction of research is advancing, but humans still need to do the hardest part: making sure these technologies actually improve learning, not just replace old complexities with new, more expensive ones.

Table 2. number of publications

Country	Number of publications
USA	2
Canada	1
Bulgaria	1
Germany	1
Brazil	1
Fiji	1
Serbia	1
Norway	1
Tanzania	1

The use of AI in schools is related to students' academic performance

The use of AI in schools does not always correlate positively with academic performance, as its impact is highly dependent on the function of AI (learning support/feedback vs assessment), the type of assignment, as well as user readiness and pedagogical design. AI can improve performance in two ways: (1) improve learning capabilities, or (2) change the way performance is measured and "seen".

Bauer et al. found that although AI adaptive feedback was expected to be more personalized, there was no significant difference in performance, and interest instead decreased, suggesting that the use of AI does not automatically improve grades or skills, especially on structured tasks (Bauer et al., 2025). Yildirim-Erbasli & Bulut highlight problems in low-stakes formative assessments, where non-effortful responding undermines the validity of results, and they propose conversation-based assessment (CBA) to increase engagement and reflect actual abilities (Yildirim-erbasli & Bulut, 2023). Tsenov & Bakracheva explain that AI attitudes and literacy affect AI use and ultimately performance, with the risk of AI becoming a shortcut if it is not properly integrated in education (Tsenov & Bakracheva, 2025).

Research trends show a shift in focus from technology as a solution to socio-technical systems influenced by attitudes and literacy (Tsenov & Bakracheva, 2025), as well as from efficiency to multidimensional evaluations that include motivation and participation (Bauer et al., 2025; Yildirim-erbasli & Bulut, 2023). Research gaps include a lack of a cross-actor framework and stronger empirical evidence in the K-12 context (Bauer et al., 2025).

Aslan et al. show that AI in blended physical–digital design can improve engagement and learning outcomes, but it needs long-term validation (Aslan et al., 2024). Jokhan et al. highlight the role of AI in predicting performance for early intervention, but this model can be biased if digital access is limited (Jokhan et al., 2022). Leite assessed GenAI in chemistry learning and found great potential, but its impact on grades has not been tested (Leite, 2024). Ilić et al. emphasize the need for more consistent and contextual evidence of performance impact in e-learning (Ilić et al., 2023).

Mwakalinga & Mabilika found that although students feel AI improves academic achievement, educators are concerned that students do not understand AI work products, so digital literacy and ethical guidelines are needed (Mwakalinga & Mabilika, 2025). Terrazas-Arellanes et al. show that AI-MLS in scoring can speed up assessment and provide insights for differentiated instruction, but requires teacher training and privacy governance (Terrazas-arellanes et al., 2025).

In conclusion, the relationship between AI and academic performance is not only measured by grades, but by student understanding and engagement. AI can improve performance through more consistent assessments and more appropriate interventions, but only if pedagogically and ethically integrated, so that results that appear to reflect real competencies (Mwakalinga & Mabilika, 2025; Terrazas-arellanes et al., 2025).

Conditions that make AI improve and decrease academic performance

The digital paradox in the use of AI in education arises when technology that is supposed to improve academic achievement does not always have a consistent positive impact. Some studies show that AI can improve performance only if there is a match between AI features, task types, and pedagogical design, while under improper conditions, AI can degrade performance by interfering with students' motivation and psychological needs, or being used without adequate digital literacy.

AI improves performance when it serves as high-quality instructional support. Bauer et al. affirm that AI adaptive feedback is effective when the system is ready and tasks are clearly defined, but the benefits depend on a mature pedagogical design (Bauer et al., 2025). Yildirim-Erbasli & Bulut show that low-stakes digital assessments can trigger non-effortful responses, reducing the validity of results, but conversation-based assessments (CBAs) can improve engagement and performance (Yildirim-erbasli & Bulut, 2023). Without proper pedagogical design, AI actually lowers interest in learning and performance, as found by Bauer et al. on structured statistical tasks (Bauer et al., 2025). Tsenov & Bakracheva emphasize that AI without critical thinking can lead to dependency and misconception (Tsenov & Bakracheva, 2025).

In elementary education, Aslan et al. show that AI is effective when designed in an immersive and multimodal way, such as in Kid Space, which improves engagement and learning outcomes (Aslan et al., 2024). However, the long-term impact needs to be tested. Jokhan et al. states that generative AI supports academic performance if verified, but can lead to misconceptions if not critically examined (Jokhan et al., 2022). Jokhan et al. show that AI can improve performance with an early warning system, although poor implementation actually increases the burden on students (Jokhan et al., 2022). Ilić et al. emphasized that AI is effective in adaptive assessment, but it is risky if students have low self-regulation (Ilić et al., 2023).

Mwakalinga & Mabilika point out that while many students feel AI improves achievement, many educators are concerned that students do not understand the results of AI's work, leading to dishonesty and a decline in critical thinking (Mwakalinga & Mabilika, 2025). Terrazas-Arellanes et al. show that AI-MLS can improve assessment consistency, but its success depends on teacher readiness and good governance (Terrazas-arellanes et al., 2025).

Synthesize, AI improves performance when adapted to pedagogy, used as a support, and balanced with literacy and strong governance (Aslan et al., 2024; Bauer et al., 2025; Yildirim-erbasli & Bulut, 2023). In contrast, AI degrades performance when applied without pedagogical sensitivity, disrupts motivation, and is managed without adequate readiness (Ilić et al., 2023; Jokhan et al., 2022; Mwakalinga & Mabilika, 2025; Tsenov & Bakracheva, 2025).

The most relevant school strategies for navigating the AI–academic performance intersection

The intersection of AI and academic performance in schools is not always linear, with the impact of AI depending on the suitability of AI for assignments, literacy, usage limits, and how schools maintain motivation to keep AI a "tool" rather than a "substitute" for thinking. Bauer et al. show that AI adaptive feedback does not necessarily improve performance, especially on structured tasks, and can decrease interest in learning (Bauer et al., 2025). Tsenov & Bakracheva emphasize the importance of AI literacy to avoid non-reflective use and ensure AI supports competencies, not just outputs (Tsenov & Bakracheva, 2025). Yildirim-Erbasli & Bulut propose conversation-based assessment (CBA) to improve the engagement and validity of assessments (Yildirim-erbasli & Bulut, 2023).

Relevant school strategies are: (1) ensuring the match of tasks with AI so as not to diminish interest (Bauer et al., 2025), (2) building literacy and mindful usage limits so that AI does not become a shortcut (Tsenov & Bakracheva, 2025), and (3) designing assessments that validate thought processes through CBA (Yildirim-erbasli & Bulut, 2023). Additionally, Jokhan et al. show that AI can serve as an early warning system for intervention (Jokhan et al., 2022), while Ilić et al. emphasize the need for AI in pedagogical design to avoid AI being just numbers (Ilić et al., 2023).

Aslan et al. added that immersive and multimodal conversational AI strengthens active engagement (Aslan et al., 2024),and Leite points out that the use of GenAI must be complemented by ethical policies (Leite, 2024). Mwakalinga & Mabilika emphasize the importance of literacy and ethics to ensure AI does not diminish academic comprehension or integrity (Mwakalinga & Mabilika, 2025), while Terrazas-Arellanes et al. point out that AI can accelerate and improve performance-based assessments (Terrazas-arellanes et al., 2025).

In conclusion, schools need to organize the relationship between technology, learning processes, and evidence of achievement so that improved academic performance reflects true competencies, not just visible outputs (Aslan et al., 2024; Ilić et al., 2023; Jokhan et al., 2022; Leite, 2024; Mwakalinga & Mabilika, 2025; Terrazas-arellanes et al., 2025).

Discussion

The digital paradox in schools—in the context of this study—can be understood as a socio-technical condition when technology (especially AI) is adopted with the aim of improving the quality of learning, but simultaneously produces the opposite impact: decreased motivation, weakened engagement, academic integrity is threatened, or inequality of access and learning outcomes is sharpened. A synthesis of the nine publications you compiled shows that this paradox is not a "minor side effect", but rather a recurring pattern that emerges when schools focus on tool adoption without strengthening pedagogical design, assessment design, and ethical governance and literacy.

On a conceptual level, the findings on attitudes towards AI make it clear that acceptance of technology is not binary (agree/disagree), but rather ambivalent and dependent on the context of use. Tsenov and Bakracheva point out that AI is often seen as helping task performance (speeding up work, making it easier to access information), but at the same time raises concerns about usage limits, social impacts, and risks associated with certain domains; therefore, they emphasize that the formation of AI competencies and mindful use should be on the educational agenda (Tsenov & Bakracheva, 2025). This ambivalence is important because it explains the "root paradox" in schools: when students see AI as a pragmatic aid, while some educators see it as a threat to the depth of thinking and learning goals. In other words, the digital paradox is not only happening because technology is "less than perfect", but because the social and psychological meaning of AI in the classroom has not been agreed upon and has not been laid out.

That paradox becomes even more concrete when AI is tested as a learning intervention. Bauer et al. tested AI-generated adaptive feedback versus static expert feedback in statistical learning. Theoretically, adaptive feedback should improve learning outcomes because it is more personalized. However, the results showed no significant difference in task performance, and what stood out was the decrease in interest in the AI adaptive feedback group (Bauer et al., 2025). These findings strengthen the argument that educational technology evaluation does not rely enough on cognitive indicators (grades/skills) alone, but must read cognitive-affective trade-offs. In the framework of digital paradoxes, AI can appear to be "neutral" to short-term achievement, but to the detriment of interest/motivation which is the engine of long-term learning (Bauer et al., 2025). Implicitly, schools that only chase scores can miss a latent danger: decreased interest and engagement can erode study perseverance, ultimately lowering academic performance more broadly.

The digital paradox also emerges strongly in the realm of assessment. Yildirim-Erbasli and Bulut explained that digitization of assessments (CBT/CAT) does promise administrative efficiency and scoring, but does not automatically guarantee the quality of student participation. In low-stakes formative assessments, students are more susceptible to showing non-effortful responding (answering original/fast), so scores become less valid as an indicator of academic ability; This condition risks producing wrong educational decisions (Yildirim-erbasli & Bulut, 2023). They offer conversation-based assessment (CBA) with conversation agents as a way to increase engagement and effort, but at the same time affirm the existence of fairness/equity issues in digital assessment—because the characteristics of digital interactions can benefit certain groups and harm others (Yildirim-erbasli & Bulut, 2023). From the point of view of the digital paradox, this is a classic picture: assessment systems are becoming more "modern and efficient", but the validity of learning data and the fairness of outcomes can deteriorate if the design is not sensitive to students' motivations and differences.

If the first three articles show a paradox in terms of attitudes, learning processes (feedback), and formative assessments, then the next four articles expand the paradox to the level of pedagogical design, e-learning ecosystem, performance prediction, and GenAI as a learning resource. Aslan et al. point out an interesting new trend: instead of "reducing technology" to respond to screen time concerns and lack of physical/social activity, they designed a technology that restores the physical-collaborative element through a multimodal immersive system with conversational agents. Their field studies in primary school contexts reported high engagement and increased learning gains, but also emphasized the need for longitudinal validation to ensure the sustainability of impact (Aslan et al., 2024). In the language of digital paradox, this approach tries to "solve the paradox" through design: technology is made so as not to sacrifice aspects of child development. But at the same time, the limitations of the duration of the study and the need for long-term evidence point to an important gap: many positive claims of AI in schools are still vulnerable to the "novelty effect" and are not yet strong enough for broad generalizations (Aslan et al., 2024).

At the level of online learning systems, Ilić et al. mapped the shift in research from just infrastructure to the application of smart techniques such as intelligent tutoring systems, learning analytics, recommendations, student modeling, and adaptive assessment. They emphasize the potential for personalization and learning support, but also underscore the need for more robust research on how these techniques affect the learning process as a whole, including motivation, attitudes, and performance (Ilić et al., 2023). These findings are intertwined with a pattern of digital paradoxes: personalization innovation is growing rapidly, while data readiness, institutional capacity, and pedagogical readiness often lag behind. As a result, systems can "look sophisticated" but do not automatically improve the quality of learning in a whole and fair manner (Ilić et al., 2023).

Jokhan et al. show another trend that is increasingly dominant: the use of AI/ML for performance prediction (early warning) so that interventions can be carried out earlier. Predictive models such as Random Forest are reportedly able to

provide early predictions by week 6 with high accuracy, so that institutions can help vulnerable students before it is too late (Jokhan et al., 2022). However, this is where the digital paradox turns into a structural paradox: the benefits of predictive AI depend on the quality and intensity of the digital footprint. If a classroom/school has low digital activity or limited access, the data needed for predictions may be inadequate, so groups that are already left behind are at risk of falling further behind in the "benefits of AI" (Jokhan et al., 2022).In other words, the more digital an ecosystem is, the more "readable" and helpful it becomes—and this holds the risk of systemic bias.

In the realm of GenAI, Leite shows that generative chatbots (e.g. ChatGPT, Gemini, Copilot) can generate coherent responses and potentially help with chemical learning, but their implementation requires critical supervision and literacy; GenAI's output needs to be read as probabilistic results that must be verified and used actively-critically, not accepted as final truth (Leite, 2024).Leite also asserts that the rate of change in GenAI's capabilities is so rapid that research evidence can quickly become obsolete, reinforcing the paradox of "technology lagging behind policy" (Leite, 2024). This is in line with other findings: schools want to take advantage of the latest innovations, but the evidence base and regulatory tools are often lagging behind the pace of AI system updates.

The last two articles show a paradox in the realm of classroom culture and AI-based grading systems. Mwakalinga and Mabilika point out that AI is already in the plural use, but the perceptions of students and educators differ sharply: students tend to see AI as improving academic achievement, while educators are more concerned about declining cognitive comprehension and engagement. They also emphasize the risk of academic dishonesty and a decline in critical thinking, thus driving the need for ethical rules, digital/AI literacy, and clear institutional norms (Mwakalinga & Mabilika, 2025). This means that the digital paradox in schools is often not a matter of "AI is right or wrong", but a matter of governance: without rules and literacy, AI drives academic output that looks good but is competively fragile.

In contrast, Terrazas-Arellanes et al. point out the "solution" side of AI in assessment: machine learning systems for scoring constructive answers on performance-based science assessments can reach substantial agreement with human assessors and have the potential to provide actionable insights for differentiated learning, while reducing the burden on teachers (Terrazas-arellanes et al., 2025). But they also emphasize that implementation demands teacher readiness, data governance, and attention to fairness and transparency—without it, AI in assessment can give rise to mistrust, new biases, or simply replace old problems with more complex ones (Terrazas-arellanes et al., 2025). Here the digital paradox emerges as a trade-off: efficiency and consistency vs governance demands and implementation competence.

If all these results are fused together, then the trend of digital paradoxical research in schools moves clearly in three directions: first, from "technology as a solution" to "technology as a socio-technical system" that is influenced by attitudes, literacy, and beliefs (Mwakalinga & Mabilika, 2025; Tsenov & Bakracheva, 2025). Second, from a single score-based evaluation to a multidimensional evaluation that includes motivation, interest, effort, and quality of participation (Bauer et al., 2025; Yildirim-erbasli & Bulut, 2023). Third, from the efficiency of digitalization to attention to fairness, equity, and structural bias due to differences in access and quality of data (Jokhan et al., 2022; Yildirim-erbasli & Bulut, 2023) . This means that the digital paradox is increasingly understood as a matter of design and governance, not a matter of "the presence/absence of technology".

The most consistent research gaps also emerge emphatically from this synthesis. First, empirical evidence for AI in learning is still often concentrated in specific contexts (e.g., higher education), so generalization to K–12 requires more contextual and, ideally, longitudinal studies (Aslan et al., 2024; Bauer et al., 2025). Second, much of GenAI's research still focuses on content quality/accuracy of user responses or perceptions, while impacts on academic integrity, critical thinking, and changes in learning behavior require stronger evaluation designs (Leite, 2024; Mwakalinga & Mabilika, 2025). Third, the issues of fairness and bias have often not been positioned as the main outcomes; Even though the digital paradox is often "invisible" if you only read the average score—inequality can increase even if the average appears stable (Jokhan et al., 2022; Yildirim-erbasli & Bulut, 2023). Fourth, intelligent engineering innovation in e-learning is growing rapidly, but the gap between technological capabilities and institutional readiness (data, human resources, pedagogy, governance) is still wide (Ilić et al., 2023; Terrazas-arellanes et al., 2025).

Thus, this discussion leads to one conceptual conclusion: the digital paradox in schools occurs when AI improves "visible performance" (assignment products, scoring efficiency, performance predictions), but does not automatically improve "actual learning quality" (understanding, motivation, effort, fairness). Navigating the AI–academic performance intersection means schools need to shift the focus from tool adoption to ecosystem design: *task–AI fit* and affective impact (Bauer et al., 2025), assessment design that maintains validity and effort (Terrazas-arellanes et al., 2025; Yildirim-erbasli & Bulut, 2023), ethical governance and AI literacy (Mwakalinga & Mabilika, 2025; Tsenov & Bakracheva, 2025), as well as mitigating structural biases arising from data and access disparities (Jokhan et al., 2022). At this point, the "successful AI" is not the most advanced AI, but rather the AI that is most aligned with pedagogical goals, the most accountable, and the most equitable for all students.

Conclusion

Based on the results of the synthesis and discussion, the main findings of this study confirm that *the digital paradox* in schools arises because AI works like a "double-edged sword": it can strengthen learning outcomes while weakening the quality of learning, depending on *how* and *in the context* of its use. In the framework of this study, the paradox is most clearly seen at the intersection of AI as *a learning support* (scaffolding, feedback, adaptive exercise, 24/7

tutor) versus AI as *a task replacement* (replacing the thought process with instant results). When AI functions as *a learning support*, it has the potential to improve academic performance through personalization, quick feedback, and reinforcement of independent learning. However, when AI shifts to *task replacement*, the risks lead to cognitive dependence, superficial understanding, misconceptions, and erosion of academic integrity. This "support vs replacement" framework also parallels the dichotomy *of ethical-directed* versus *pragmatic-instant* use that makes the impact of AI can be opposite in different school contexts.

The next findings emphasize that the AI–academic performance relationship is never "single" and is not automatically positive. The empirical evidence in the literature also tends to be varied/inconsistent—some show improved outcomes, some are neutral, and some highlight negative consequences—which means that the effects of AI are highly sensitive to institutional readiness, educator competence, student motivation, assignment/assessment design, and AI access and literacy. Therefore, the visible academic performance (grades, quality of assignments) needs to be read together with process indicators (involvement, depth of reasoning, validity of effort during assessment) so that schools do not get stuck in "pseudo-achievements" that are neat on the surface but fragile in understanding.

Finally, the main finding that is the "locking" of the discussion is the need to navigate the AI–academic performance intersection through explicit maps of variables: mediators (e.g., self-regulated learning, quality of understanding, engagement, cognitive load) and moderators (e.g., AI literacy, teacher support/supervision, school policies and culture, types of HOTS vs. routine assignments, and technology access). This map is important because it is where the digital paradox can be explained operationally: when AI is a lever for academic performance, and when it is a source of declining learning quality.

Recommendations

In order for schools to leverage AI to improve academic performance without creating new paradoxes, AI must be positioned as part of the learning ecosystem, not just an application. Schools need to map out the use of AI that actually supports learning (concept exploration, adaptive exercise, formative feedback) and limit AI to tasks that easily turn into substitutes for thinking (end product without reasoning).

To be effective, schools are required to build AI literacy and ethics as core competencies: getting used to verification (checking sources, consistency, logic), understanding biases/limitations, and reflecting on right-wrong reasons, so that dependencies, misconceptions, or integrity violations do not occur. At the same time, assessments need to be redesigned to assess the thought process (steps, reasons, reflections, revisions), with AI being able to help with initial scoring/feedback while maintaining teacher agency and fairness.

All of this needs to be supported by teacher training and clear data governance (privacy, access, transparency) so that AI does not become a new burden or source of inequality. With pedagogical signs, reasoning-based assessments, ethical literacy, and fair governance, academic performance increases because competence grows, not just the output gets neater.

References

Adewale, M., Azeta, A., Abayomi-Alli, A., & Sambo-Magaji, A. (2024). Impact of artificial intelligence adoption on students' academic performance in open and distance learning: A systematic literature review. *Heliyon*, *10*. https://doi.org/10.1016/j.heliyon.2024.e40025

Aslam, F., Aslam, F., Marwat, S. A., Arif, M., Khadim, A., & Hussain. (2025). Artificial Intelligence in Higher Education: Shaping the Future of University Teaching Through Adaptive Learning, Intelligent Tutoring, and Academic Analytics. *Indus Journal of Social Sciences*. https://doi.org/10.59075/ijss.v3i2.1007

Aslan, S., Durham, L. M., Alyuz, N., Okur, E., Sharma, S., & Savur, C. (2024). Computers and Education : Artificial Intelligence Immersive multi-modal pedagogical conversational artificial intelligence for early childhood education : An exploratory case study in the wild. *Computers and Education: Artificial Intelligence*, *6*(November 2023), 100220. https://doi.org/10.1016/j.caeai.2024.100220

Bahroun, Z., Anane, C., Ahmed, V., & Zacca, A. (2023). Transforming Education: A Comprehensive Review of Generative Artificial Intelligence in Educational Settings through Bibliometric and Content Analysis. *Sustainability*. https://doi.org/10.3390/su151712983

Bauer, E., Richters, C., Pickal, A. J., Klippert, M., Sailer, M., & Stadler, M. (2025). *Effects of AI- - generated adaptive feedback on statistical skills and interest in statistics : A field experiment in higher education*. *November 2024*, 1735–1757. https://doi.org/10.1111/bjet.13609

Dogan, M. E., Dogan, T. G., & Bozkurt, A. (2023). The Use of Artificial Intelligence (AI) in Online Learning and Distance Education Processes: A Systematic Review of Empirical Studies. *Applied Sciences*. https://doi.org/10.3390/app13053056

Francis, N., Jones, S., & Smith, D. (2025). Generative AI in Higher Education: Balancing Innovation and Integrity. *British Journal of Biomedical Science*, *81*. https://doi.org/10.3389/bjbs.2024.14048

Gligorea, I., Cioca, M., Oancea, R., Gorski, A., Gorski, H., & Tudorache, P. (2023). Adaptive Learning Using Artificial Intelligence in e-Learning: A Literature Review. *Education Sciences*. https://doi.org/10.3390/educsci13121216

Gustilo, L., Ong, E., & Lapinid, M. (2024). Algorithmically-driven writing and academic integrity: exploring educators' practices, perceptions, and policies in AI era. *International Journal for Educational Integrity*, *20*, 1–43. https://doi.org/10.1007/s40979-024-00153-8

Hamal, O., Faddouli, N.-E. El, Harouni, M. H. A., & Lu, J. (2022). Artificial Intelligent in Education. *Sustainability*. https://doi.org/10.3390/su14052862

Ilić, M., Mikić, V., Kopanja, L., & Vesin, B. (2023). Intelligent techniques in e - learning : a literature review. In *Artificial Intelligence Review* (Vol. 56, Issue 12). Springer Netherlands. https://doi.org/10.1007/s10462-023-10508-1

Jokhan, A., Chand, A. A., & Singh, V. (2022). *Increased Digital Resource Consumption in Higher Educational Institutions and the Artificial Intelligence Role in Informing Decisions Related to Student Performance*.

Kamalov, F., Calonge, D. S., & Gurrib, I. (2023). New Era of Artificial Intelligence in Education: Towards a Sustainable Multifaceted Revolution. *Sustainability*. https://doi.org/10.3390/su151612451

Khan, M., & Irfan, R. (2025). The Role of Artificial Intelligence in Academic Achievement in the Current Scenario. *International Journal For Multidisciplinary Research*. https://doi.org/10.36948/ijfmr.2025.v07i04.49965

Leite, B. S. (2024). *Journal of Applied Learning & Teaching content responses*. *7*(2).

Majeed, D., Khan, A. Y., Munir, M., Shan, R. U., Khalid, H., Khan, K., Aurangzeb, T., & Atique, E. (2025). Opportunities and Challenges of Generative AI in Pakistani Higher Education: A Qualitative Study on Student Perspectives in Learning, Integrity, and Innovation. *Nature-Nurture Journal of Psychology*. https://doi.org/10.53107/nnjp.v4i2.89

Mariñoso, P. E., García, P. Á. C., Jareño, I. M., & Rodriguez, E. C. (2025). Evaluating the use of AI in the design of learning situations by university students of early childhood education. *Frontiers in Psychology*, *16*. https://doi.org/10.3389/fpsyg.2025.1604414

Mwakalinga, S. E., & Mabilika, F. A. (2025). Perceptions , pitfalls , and proposals for the ethical use of artificial intelligence in the classroom : a case study of students and educators (teachers ' and lecturers '). *Cogent Education*, *12*(1). https://doi.org/10.1080/2331186X.2025.2557611

Perkins, M., Furze, L., Roe, J., Vietnam, J. M. B. U., University, D., & Singapore, J. C. U. (2023). The Artificial Intelligence Assessment Scale (AIAS): A Framework for Ethical Integration of Generative AI in Educational Assessment. *Journal of University Teaching and Learning Practice*. https://doi.org/10.53761/q3azde36

R, J., Arif, H., U, H. B., T, R., Hanumanthappa, H., & S, G. R. (2025). Personalized Learning Systems Using AI Enhancing Adaptive Education. *International Journal of Environmental Sciences*. https://doi.org/10.64252/ps80c806

Strielkowski, W., Grebennikova, V., Lisovskiy, A., Rakhimova, G., & Vasileva, T. (2024). AI-driven adaptive learning for sustainable educational transformation. *Sustainable Development*. https://doi.org/10.1002/sd.3221

Terrazas-arellanes, F. E., Strycker, L., Alvez, G. G., Miller, B., & Vargas, K. (2025). *Promoting Agency Among Upper Elementary School Teachers and Students with an Artificial Intelligence Machine Learning System to Score Performance-Based Science Assessments*.

Tsenov, M. Y., & Bakracheva, M. A. S. (2025). Attitudes towards artificial intelligence in professional and personal life. *The Education and Science Journal*, *27*(2), 159–174. https://doi.org/10.17853/1994-5639-2025-2-159-174

Yarlagadda, K. C. (2025). AI in Education: Personalized Learning and Intelligent Tutoring Systems. *European Journal of Computer Science and Information Technology*. https://doi.org/10.37745/ejcsit.2013/vol13n321527

Yildirim-erbasli, S. N., & Bulut, O. (2023). Computers and Education : Artificial Intelligence Conversation-based assessment : A novel approach to boosting test-taking effort in digital formative assessment. *Computers and Education: Artificial Intelligence*, *4*(June 2022), 100135. https://doi.org/10.1016/j.caeai.2023.100135

Zhai, C., Wibowo, S., & Li, L. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review. *Smart Learning Environments*, *11*. https://doi.org/10.1186/s40561-024-00316-7