



OPEN ACCESS

EDITED BY

Sami Heikkinen,
LAB University of Applied Sciences, Finland

REVIEWED BY

Ijeoma John-Adubasim,
United Kingdom
Omaia Al-Omari,
Prince Sultan University, Saudi Arabia

*CORRESPONDENCE

Latifah Hamdan Alghamdi
✉ lalghamdi@kku.edu.sa

RECEIVED 13 September 2025

ACCEPTED 16 October 2025

PUBLISHED 13 November 2025

CITATION

Alghamdi LH and Alghizzi TM (2025)
Educators' reflections on AI-automated
feedback in higher education: a structured
integrative review of potentials, pitfalls, and
ethical dimensions.
Front. Educ. 10:1704820.
doi: 10.3389/feduc.2025.1704820

COPYRIGHT

© 2025 Alghamdi and Alghizzi. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Educators' reflections on AI-automated feedback in higher education: a structured integrative review of potentials, pitfalls, and ethical dimensions

Latifah Hamdan Alghamdi^{1*} and Talal Musaed Alghizzi²

¹King Khalid University, Abha, Saudi Arabia, ²Department of English Language and Literature, College of Languages and Translation, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh, Saudi Arabia

The rapid incorporation of artificial intelligence (AI) into higher education has established automated feedback systems as both a potential benefit and a challenge. Accordingly, this systematic study synthesizes the findings of 37 empirical investigations (2014–2024) to underscore the significance of teachers' perspectives, which are sometimes overlooked in the use of AI-mediated feedback. Research indicates that AI can enhance customization, deliver immediate feedback, optimize repetitive processes, and increase student engagement. Nonetheless, these advantages are persistently compromised by concerns regarding algorithmic bias, data privacy, the deterioration of teacher-student relationships, and inadequate professional growth. The current evidence base is methodologically deficient, predominantly including short-term research or subjective evaluations, with just a limited number providing longitudinal data or controlled comparisons. This research distinguishes itself from previous evaluations that emphasize technology attributes or student results by integrating the FATE framework (Fairness, Accountability, Transparency, Ethics) with adoption models (TAM/UTAUT). It redefines educators as proactive mediators whose ethical choices and professional identities influence the optimal integration of AI. Thus, it contends that AI feedback should enhance, rather than replace, human teaching, and that its ongoing application depends on professional growth and strong governance frameworks. Future research must focus on longitudinal, cross-cultural, and outcome-validated approaches to shift the profession from experimental excitement to evidence-based educational change.

KEYWORDS

AI-powered feedback, higher education, teacher perspectives, ethics, personalization, adoption frameworks

1 Introduction

The rapid emergence of artificial intelligence (AI) in higher education has generated both enthusiasm and concern among educators, politicians, and academics. Recent research has underscored the potential of AI-driven feedback systems to provide responses that are not only immediate but also customized and adaptable to the diverse needs of students (Bond et al., 2024; Samarescu et al., 2024). Harnessing the power of machine learning algorithms, these systems are capable of analyzing performance data, pinpointing specific areas that require improvement, and offering tailored recommendations. Such functions stand in marked contrast to the slower and more rigid practices of traditional grading methods. As Holmes

et al. (2019) argue, these capabilities significantly improve scalability and timeliness, while simultaneously fostering a more personalized learning experience that complements, rather than replaces, the essential role of human engagement.

The integration of AI into educational environments offers both advantages and significant challenges. Educators and researchers persist in voicing apprehensions over data privacy, algorithmic prejudice, and ethical responsibilities, with concerns that AI may increasingly encroach upon human feedback and undermine the relational dimension of education (Kim and Kim, 2022). The autonomy, willingness, and preparedness of educators to implement AI technology are essential criteria for its impact on educational practices. Insufficient professional development and weak ethical protections in education may lead to the use of AI that exacerbates existing inequalities and undermines the human-centered values of the field (Crompton and Burke, 2023).

This argument is supported by evaluations that highlight both benefits and disadvantages. Artificial intelligence has continuously demonstrated potential in reducing administrative burdens, enhancing engagement, and delivering customized learning trajectories (Chounta et al., 2021; Fullan et al., 2023). A considerable percentage of the current data relies on self-reported perceptions, transient experimental projects, or narrowly defined groups, instead of robust, outcome-focused research. Consequently, data often display inconsistency, particularly in illustrating lasting effects on learning outcomes (Zawacki-Richter et al., 2019; Memarian and Doleck, 2023). This discrepancy between optimism and empirical evidence underscores the pressing need for systematic assessments that employ extensive data and objectively evaluate both the benefits and drawbacks of AI-mediated feedback.

Previous reviews, such as those by Zawacki-Richter et al. (2019) and Bond et al. (2024), laid important groundwork by cataloging AI applications, technical affordances, and student outcomes; they paid relatively little attention to the professional and ethical dimensions of teachers' roles in mediating AI adoption. The ethical and professional responsibilities of educators in the implementation of AI have, however, been overlooked in previous contributions. This study redefines teachers as active mediators instead of passive end-users, highlighting that their judgements and decisions play a crucial role in the outcomes of technological deployment. In doing so, it employs adoption models such as TAM and UTAUT in conjunction with ethical frameworks like FATE, thereby situating AI integration within the wider context of professional identity, trust, and moral responsibility.

This study redirects attention from prior research that mainly investigated technological effectiveness or student outcomes to highlight instructors' perspectives, concerns, and professional experiences with AI-generated feedback. This synthesis encompasses 37 peer-reviewed studies published from 2014 to 2024, classified into three interconnected themes: first, educators' perceptions and use of AI feedback tools; second, the identified benefits and challenges; and third, the degree to which ethical and pedagogical considerations impact the adoption process.

The subsequent research questions (RQs) were established to direct the review:

RQ1: What are educators' overall perceptions of AI-generated feedback in higher education?

RQ2: What distinct advantages do educators identify in AI-driven feedback systems?

RQ3: What apprehensions do educators express concerning the incorporation of AI feedback systems?

RQ4: What obstacles and circumstances affect the successful implementation of AI in educational practices?

This study seeks to address these questions, facilitating a balanced, evidence-based synthesis that informs both academia and practice. To address the need for a more innovative problem statement, this study reframes the issue of AI-automated feedback not merely as a technological innovation but as a pedagogical and ethical transformation mediated by teachers. Existing research has largely emphasized technical efficiency and student outcomes, overlooking the dual dimensions of adoption and accountability that underpin effective integration. Drawing on the Technology Acceptance Model (TAM) and the Unified Theory of Acceptance and Use of Technology (UTAUT), this review situates educators as active agents whose perceptions of usefulness, ease of use, and institutional support determine the success of AI adoption. At the same time, ethical concerns are embedded through the FATE framework (Fairness, Accountability, Transparency, and Ethics), highlighting that issues of trust, fairness, and professional autonomy condition teachers' engagement with AI. By merging these frameworks, this study advances an integrative problem statement: that the effective use of AI-mediated feedback in higher education depends not only on its technical functionality but also on how teachers ethically and pedagogically negotiate their roles within digitalized feedback ecosystems.

2 Literature review

2.1 Teachers' perspectives on AI feedback

Teachers tend to approach AI-driven feedback with a tempered sense of optimism, recognising both its potential and its limitations. On the one hand, educators highlight its capacity to enhance individualised instruction, streamline productivity by automating repetitive tasks, and create more adaptable learning pathways for students (Karaca and Kilcan, 2023). The immediacy and comprehensiveness of AI-generated feedback are particularly valued, as they allow learners to address gaps in understanding while the material remains cognitively relevant (Fullan et al., 2023). In addition, such tools are seen as valuable aids in lesson design, offering diagnostic insights and tailoring information to accommodate the needs of diverse learners (Roll and Wylie, 2016; Chounta et al., 2021). This cautious optimism is, however, tempered by ongoing misgivings. Many educators fear that the use of AI tools may erode the human dimension of teaching, especially the relational and emotional bonds that are essential to meaningful student engagement and learning (Rosak-Szyrocka et al., 2023).

Educators express enthusiasm for the educational benefits of AI, asserting that these tools should augment rather than supplant human instruction. This viewpoint highlights valuable insights from research on technology adoption, showing that successful implementation

thrives on factors like perceived benefit, usability, and system reliability. In practice, educators' readiness to embrace AI-driven feedback is influenced by ethical precautions, institutional support, congruence with pedagogical principles, and the operational efficacy of the tools themselves.

2.2 Recent advancements in AI-driven feedback

Over the past decade, an exciting surge has been observed in the study of artificial intelligence in education, particularly since 2018 with the advent of adaptive learning platforms and generative AI (Wu et al., 2023). Innovations such as intelligent tutoring systems, chatbots, and automated assessment tools have been introduced with the aim of easing educators' cognitive load and personalising student learning (Kim and Kim, 2022). Educators often appreciate these advances for their ability to improve student engagement and facilitate real-time formative assessment (Chounta et al., 2021; Fullan et al., 2023). Adaptive platforms dynamically adjust the complexity of content to ensure that advanced learners are consistently challenged, while immediate support is provided to struggling students. Similarly, collaborative AI tools have been acknowledged for enhancing peer-to-peer learning and motivation (Yang and Kyun, 2022).

The literature on AI in education is ambiguous and contentious. Recent studies demonstrate that AI tools can increase student engagement and productivity, but they also pose risks such as distraction, oversimplified tasks, and unequal access to resources (Crompton and Burke, 2023). Furthermore, improving the clarity of algorithmic decision-making presents a significant opportunity to enhance instructors' credibility, alleviate transparency concerns, and promote greater engagement with AI tools (Benneh, 2023). Research consistently demonstrates that educator involvement and ongoing professional development are crucial for tackling these issues, with established methods available to mitigate such constraints. Targeted training empowers educators to enhance their skills in critically assessing and contextualizing automated outputs, effectively addressing the challenges posed by the opaque characteristics of "black box" technology (Memarian and Doleck, 2023). Recent evidence suggests that the effectiveness of AI-driven feedback is shaped primarily by its congruence with educational objectives, the safeguarding of teacher autonomy, and the preparedness of institutions for integration, rather than by the sophistication of the technology alone.

2.3 Synthesis and research gaps

2.3.1 Gaps emerge across the literature

- **Methodological limitations:** A substantial portion of contemporary research mostly relies on self-reported data or short-term interventions, which, while beneficial for generating preliminary insights, provide a limited basis for assessing long-term effectiveness.
- **Inadequate theoretical integration:** Although adoption frameworks (e.g., TAM, UTAUT) and pedagogical theories are pertinent, there exists a compelling opportunity for research to

more explicitly integrate teachers' viewpoints within these models, thereby enhancing conceptual underpinnings.

- **Ethical and relational dimensions:** Although fairness, privacy, and transparency are acknowledged concerns, there is a considerable scope for more empirical research to explore how educators are creatively tackling these challenges in practice.

This review highlights teacher perspectives as essential elements that could significantly enhance educational equity, engagement, and trust through AI feedback, rather than only perceiving them as attitudes towards technology.

2.3.2 Positioning against prior reviews

Earlier syntheses, most notably Zawacki-Richter et al. (2019) and Bond et al. (2024), laid important groundwork but differed markedly in focus. Zawacki-Richter et al. (2019) structured their review around categories of AI applications (e.g., tutoring, grading, analytics) and student outcomes, with educators treated primarily as end-users rather than ethical mediators. Bond et al. (2024) advanced the field by calling for more attention to ethics and collaboration, yet their analysis emphasized system-level safeguards (bias mitigation, transparency, governance) rather than embedding ethics in teachers' professional responsibilities.

In contrast, the present review explicitly codes for teacher development, professional identity, and ethical mediation as central categories of analysis. By integrating TAM/UTAUT adoption models with the FATE framework, this review bridges adoption behavior with ethical accountability, positioning teachers as gatekeepers of trust and pedagogical integrity. This dual-lens analysis represents a conceptual advance beyond prior reviews, which largely treated ethics as external to educators and adoption as a matter of technical utility.

3 Methods

3.1 Inclusion criteria and screening

This study employed a systematic strategy to locate pertinent literature, culminating in the selection of 37 empirical studies that met the established inclusion criteria (refer to Table 1). Research must investigate aspects of AI development in education, include a clearly stated methodology section, and concentrate on AI applications inside formal educational environments to qualify for inclusion. The research excluded from the analysis failed to meet these criteria, particularly those focused on informal learning contexts or demonstrating insufficient methodological transparency.

TABLE 1 Overview of the selection criteria for inclusion.

Concepts	Inclusion criteria
AI application	Must be focused on AI applications in formal educational settings
Methodology	Must have a clearly defined methodology section
Educational phase	Emphasis on preceding and forthcoming phases of AI evolution in education.
Publication date	Studies published from 2014 to 2024
Language	English-language publications

The review was restricted to English-language publications. This decision was made for two reasons: (a) English remains the dominant language of publication in international AI-in-education scholarship, ensuring access to the majority of influential works; and (b) resource constraints prevented systematic translation and coding across multiple languages. The screening process commenced with an initial assessment of titles and abstracts, leading to the exclusion of non-relevant studies or those not meeting the inclusion criteria. Subsequent to the initial screening, full-text reviews were performed on the remaining studies. Two independent reviewers evaluated each publication for relevance, methodological rigor, and its contribution to the understanding of AI-automated feedback systems from the educators' perspective. Discrepancies among reviewers were addressed through consensus discussions, and a third reviewer was consulted when necessary to achieve a final decision on inclusion.

The calculation of inter-rater reliability employed Cohen's kappa, resulting in a value of 0.91, signifying a substantial level of agreement among reviewers. The notable consistency enhances the credibility of the study selection process, diminishes the potential for bias, and improves the overall validity of the review.

3.1.1 Inter-rater reliability

Ensuring the rigor and trustworthiness of qualitative data analysis necessitates assessing inter-rater reliability. Prior to coding, all raters participated in training to familiarize themselves with the coding framework (including definitions and examples for each code). This training ensured a consistent understanding and application of criteria through guidelines, examples, and mock assessments. At least two reviewers independently evaluated each study using standardized forms. A 10% subset of the data was used for a pilot coding exercise, where each rater coded the subset independently. The results were then compared and discussed among the raters. The Cohen's kappa coefficient, calculated from this pilot coding, was 0.91, indicating almost perfect agreement between the two primary reviewers. Discrepancies in reviewers' assessments were resolved through detailed discussions and consensus; if agreement could not be reached, a third reviewer acted as an arbitrator. The elevated Cohen's kappa signifies that raters were highly consistent in applying the coding framework, thereby minimizing subjectivity and bias. This reliability is paramount for the validity of the study's findings and reinforces the trustworthiness of the conclusions drawn from the data.

3.2 Search strategy

The systematic review employed a rigorous search strategy to ensure a comprehensive collection of relevant literature. The search focused on studies published between 2014 and 2024, reflecting the rapid evolution and increased adoption of AI technologies in educational settings. Searches were conducted across Scopus, Web of Science, and Google Scholar using Boolean combinations of key terms ('AI-mediated feedback' AND 'higher education' AND 'teacher adoption'). The search was limited to peer-reviewed journal articles and conference proceedings published in English. Use-case inclusion was restricted to empirical studies involving feedback practices in higher education settings. Databases such as Scopus, Web of Science, and Google Scholar were utilized to locate peer-reviewed articles, conference papers, and dissertations. Keywords included "AI feedback

in education," "automated feedback systems," "teacher perspectives on AI," "AI in higher education," and "personalized learning through AI." Boolean operators (AND, OR, NOT) and filters helped refine the search to studies centered specifically on AI applications within formal educational contexts (as opposed to informal learning environments or purely technical AI developments). Reference lists of key articles were also reviewed to identify additional studies missed in the database search. The search was limited to English-language publications for consistency and to capture studies most likely to influence the global discourse on AI integration in higher education. Studies lacking a transparent methodology or offering only theoretical perspectives without empirical data were excluded in order to prioritize evidence-based research. This multi-pronged search strategy aimed to capture a diverse range of teacher perspectives on the efficacy, challenges, and future directions of AI-automated feedback in education.

The exact Boolean search strings used were as follows:

- Scopus: ("artificial intelligence" OR "AI") AND ("feedback" OR "automated feedback" OR "formative feedback") AND ("higher education" OR "university" OR "college" OR "tertiary education" OR "teacher")
- Web of Science: TS = ("AI" OR "artificial intelligence") AND TS = ("feedback" OR "automated feedback") AND TS = ("higher education" OR "teacher perspectives")
- Google Scholar (supplementary): "AI feedback in education" OR "AI automated feedback" OR "teacher perspectives AI feedback."

Searches were restricted to 2014–2024 to capture the period of rapid growth in AI integration into educational practice, particularly the post-2018 expansion of adaptive and generative systems. Reference lists of key studies were hand-searched to capture additional works not indexed in databases.

3.3 Data extraction and synthesis

While the review primarily employed qualitative content analysis to synthesize teachers' perspectives, additional quantitative synthesis was conducted to explore trends in the literature. For instance, the distribution of studies over the 2014–2024 period was visualized in a timeline, revealing a notable increase in publications after 2018, coinciding with the growing prominence of AI technologies in educational discourse. Furthermore, quantitative data from the included studies were compiled to provide a clearer picture of the overall impact of AI-powered feedback systems. Studies that provided statistical evidence of AI's effectiveness in enhancing personalized learning, improving student engagement, or reducing teacher workload were compared. The review highlighted that a majority of teachers in the reviewed studies expressed positive attitudes toward AI's role in personalizing instruction (e.g., Chounta et al., 2021; Kim and Kim, 2022). In a classroom-based study, Mollick and Mollick (2022) found that AI chatbots provided immediate feedback, enhancing student engagement and improving learning outcomes. Through this synthesis, the review underscored both the promise of AI-automated feedback systems and the gaps in the current evidence base. Many studies reported positive outcomes, yet the overall robustness of

evidence remains limited; several works call for more rigorous research designs and larger samples to substantiate the long-term impact of AI in education.

3.4 Study selection (PRISMA flow)

Following the PRISMA 2020 guidelines, a total of 1,246 records were initially identified across databases (Scopus = 532; Web of Science = 428; Google Scholar = 286). After removing 317 duplicates, 929 records remained for title and abstract screening. Of these, 701 records were excluded for not addressing AI-mediated feedback in education or for failing to meet the predefined inclusion criteria, leaving 228 full-text articles assessed for eligibility.

A further 191 full-text studies were excluded for the following reasons: (i) absence of a pedagogical or feedback-related component ($n = 62$); (ii) non-peer-reviewed or opinion-based sources ($n = 47$); (iii) lack of methodological transparency ($n = 53$); and (iv) focus outside higher-education or professional-training contexts ($n = 29$).

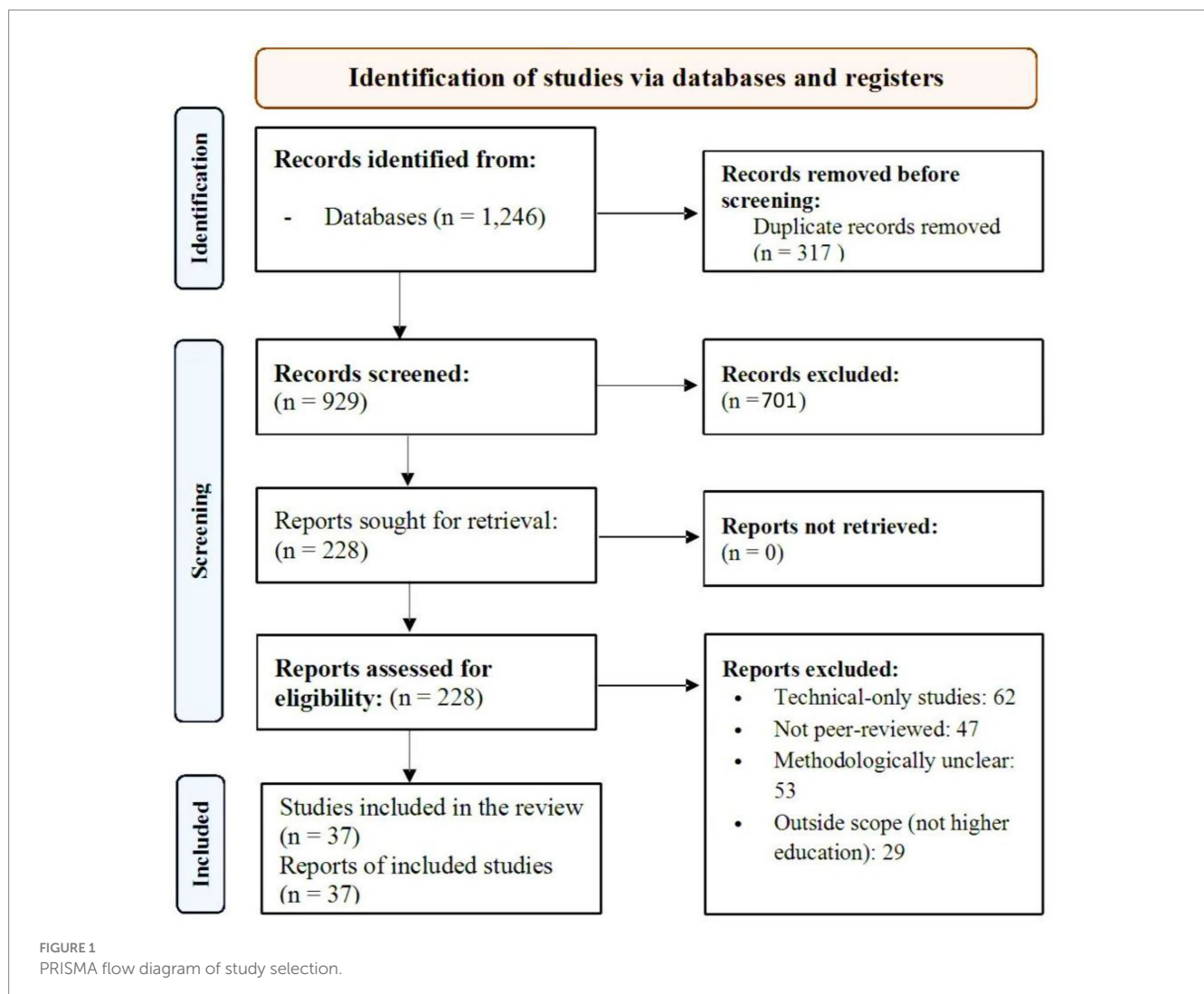
Consequently, 37 studies met all inclusion criteria and were incorporated into the final synthesis. The selection process is

summarized in [Figure 1](#), and the detailed list of included studies is presented in [Appendix A](#), which also conforms to the PRISMA 2020 checklist.

3.5 Coding schema for analysis

While the review prioritized empirical studies (quantitative, qualitative, and mixed methods), a small number of conceptual and theoretical contributions were also included. These were retained when they offered significant insights into ethical frameworks, governance structures, or teacher roles that shaped the interpretation of empirical evidence. To mitigate dilution of evidence strength, conceptual/theoretical pieces were clearly coded as non-empirical and analyzed separately from empirical findings. Their function in the synthesis was interpretive and contextual rather than evidentiary, providing a theoretical lens (e.g., TAM, UTAUT, FATE) through which the empirical data were situated.

To enable systematic comparison across studies, a coding framework was developed that categorizes the included research along five primary dimensions: context, AI tool type, study design, thematic focus, and reported outcomes.



- Context: Higher education (HE), K–12, professional development (PD), or MOOCs.
- AI tool type: Generative AI (GAI), Chatbots (CHAT), intelligent tutoring systems (ITS), feedback systems (FEED), analytics/dashboards (ANALYT), or general AI applications (GEN).
- Study design: quantitative (QUAN), qualitative (QUAL), mixed-methods (MIX), review (REV), or theoretical/conceptual (THEO).
- Thematic focus: Pedagogical effectiveness (PED), ethics and governance (ETH), teacher development and roles (TEACH), student experience (STUD), or historical/disciplinary evolution (HIST).
- Outcomes: Positive (POS), mixed (MIXED), negative (NEG), or neutral/descriptive (NEUT).

This schema not only standardized the coding process but also provided a lens for meta-level interpretation. For instance, most higher education studies employing feedback systems or learning analytics (Afzaal et al., 2021; Hooda et al., 2022) reported positive learning outcomes, whereas studies on chatbots tended to show mixed results, balancing efficiency gains with concerns over trust and a preference for human input (Escalante et al., 2023; Labadze et al., 2023). Likewise, conceptual and ethical analyses frequently leaned toward critical or cautionary conclusions (Selwyn, 2019; Benneh, 2023), reflecting the unresolved challenges of AI governance in education.

In addition to achieving high inter-rater reliability (Cohen's kappa = 0.91), the coding process included explicit adjudication procedures. When disagreements occurred, reviewers compared rationales for code assignment, consulted coding definitions, and documented the resolution pathway. Cases that could not be resolved bilaterally were escalated to a third reviewer, who justified an independent decision. All adjudicated cases were logged, creating an audit trail that reinforced transparency and minimized interpretive bias. This ensured that the synthesis of teacher perspectives was not only reliable but also traceable and reproducible.

3.5.1 Explicit coding framework

The following Table 2 illustrates a coding framework that can be summarized as follows, with illustrative examples from the literature:

3.6 Quality appraisal of evidence

To assess the robustness of the included studies, we classified them by methodological strength. Table 3 presents the methodological distribution of the eight major themes across the included studies ($N = 37$). Clear patterns emerge when themes are cross-tabulated against methodological strength. Benefits such as personalization (Tier C = 12; Tier B = 7) and immediate feedback (Tier C = 10; Tier B = 9) are strongly supported by small-scale experimental and controlled studies, suggesting that these outcomes are most visible in short-term interventions where AI-driven feedback can be isolated as an effect. Similarly, productivity gains and student engagement are consistently reported across Tiers B and C, although engagement also appears in perception-based Tier D studies ($n = 10$), indicating that it is often observed as a matter of teacher or student perception rather than validated longitudinally.

TABLE 2 Categories and coding framework with examples.

Category	Code	Definition/criteria
Context	HE	Higher education (colleges, universities)
	K12	Primary/secondary education
	PD	Professional development/teacher training
	MOOC	Massive open online courses/online-only
AI tool type	GAI	Generative AI (e.g., ChatGPT, text-to-image)
	CHAT	Chatbots
	ITS	Intelligent tutoring systems
	FEED	AI-generated feedback systems
	ANALYT	Learning analytics/dashboards
	GEN	General AI applications (not tool-specific)
Study design	QUAN	Quantitative (experiments, surveys, ML models)
	QUAL	Qualitative (interviews, case studies)
	MIX	Mixed-methods
	REV	Review (systematic, meta-analysis, narrative)
	THEO	Theoretical/conceptual analysis
Focus/theme	PED	Pedagogical effectiveness (learning outcomes, engagement)
	ETH	Ethics, governance, accountability
	TEACH	Teacher development, perceptions, and roles
	STUD	Student experience and attitudes
	HIST	Historical/disciplinary evolution
Outcomes	POS	Positive learning impact (performance, engagement, efficacy)
	MIXED	Mixed outcomes (benefits and concerns)
	NEG	Negative or critical stance
	NEUT	Neutral/descriptive (no clear evaluative outcome)

By contrast, ethical concerns, including privacy ($n = 12$ in Tier D), algorithmic bias ($n = 11$ in Tier D), and relational erosion ($n = 12$ in Tier D), cluster heavily in perception-driven designs, highlighting how these risks are often voiced as experiential or anticipatory rather than systematically tested. Interestingly, relational erosion also appears in Tier A longitudinal studies ($n = 5$), suggesting that when evidence is tracked over time, the risk to teacher–student dynamics becomes more visible. Finally, lack of training was spread across tiers but appeared most prominently in weaker designs (Tier D = 11), with only minimal attention in stronger Tier B studies ($n = 2$), reflecting a methodological imbalance in how this issue is investigated. Survey-based approaches dominate (16 out of 37 studies); notably, 12 of those 16 were conducted in university settings (i.e., three-quarters of all survey-based studies were in higher education).

Taken together, this distribution demonstrates that benefits are disproportionately validated in Tier B/C outcome-oriented designs. In contrast, concerns are more often reported in Tier D perception-based studies, with only a small number of Tier A investigations addressing them over the long term. This asymmetry highlights the reviewers' critique that enthusiasm for AI-powered feedback is primarily based on exploratory or perception-based evidence, while rigorous, longitudinal validation remains limited.

TABLE 3 Methodological characteristics of the included studies (N = 37).

Theme	Tier A (longitudinal)	Tier B (controlled)	Tier C (mixed/ experimental)	Tier D (perception)
Personalization	4	7	12	8
Immediate feedback	4	9	10	8
Productivity gains	2	7	11	6
Student Engagement	3	5	11	10
Privacy concerns	2	6	6	12
Algorithmic bias	2	4	5	11
Relational erosion	5	4	6	12
Lack of training	2	2	8	11

To further examine the distribution of evidence, the study design was cross-tabulated with educational context (Figure 2). The visualization highlights that survey-based studies dominate higher education, while experimental and longitudinal designs remain scarce in K–12 and professional development.

The cross-tabulation of study design and educational context (Table 4) reveals a strong concentration of research within higher education, where 30 of the 37 studies were conducted. Survey-based approaches dominate across all contexts (16 studies), with nearly three-quarters situated in university settings. Experimental and controlled studies remain limited ($n = 6$), and only one of these extends into K–12 education. Longitudinal and mixed-methods designs are even less common ($n = 4$), highlighting the scarcity of outcome-validated evidence. By contrast, conceptual/theoretical contributions ($n = 7$) provide interpretive frameworks but lack empirical grounding, and reviews ($n = 4$) add only partial consolidation.

4 Results

4.1 Positive attitudes and benefits of AI feedback

The synthesis of 37 studies reveals a predominantly positive outlook on AI-powered feedback among educators. Teachers highlighted four main benefits of AI feedback systems: (a) enhanced personalized instruction, (b) immediate, tailored feedback, (c) productivity gains through automation of routine tasks, and (d) enhanced student engagement (see Figure 3).

Personalized instruction was consistently the most emphasized benefit, with teachers valuing AI’s ability to adapt content to diverse student needs and provide differentiated support. Immediate feedback was also recognized as a transformative feature, allowing learners to correct errors while concepts remained cognitively salient. In terms of productivity, AI can reduce administrative burdens such as grading and attendance tracking, enabling educators to devote more time to instructional design and one-on-one support. Enhanced student engagement was attributed to AI’s interactive and adaptive features, including elements like gamification and personalized learning pathways. Where percentages are reported, they are drawn directly from individual studies and should not be interpreted as generalizable figures. Where original studies did not provide numerical data, findings are described qualitatively.

4.1.1 Immediate and tailored feedback

AI systems provide real-time, individualized feedback that traditional grading cannot match. Immediate corrections help students revise work while concepts remain fresh, reinforcing learning and improving performance (Ma et al., 2014). Tools can highlight errors in grammar, reasoning, or structure, and offer targeted suggestions for improvement. For teachers, AI’s longitudinal tracking reveals recurring difficulties and supports differentiated instruction (Chen et al., 2020; Doroudi, 2022). Overall, real-time and tailored feedback enhances student engagement and enables more responsive teaching practices.

4.1.2 Increased productivity

By automating routine tasks such as grading, attendance, and scheduling, AI reduces teacher workload, allowing for a greater focus on instruction. Automated grading provides instant results and consistent evaluation even in large classes (Benneh, 2023). Administrative tools, including AI-powered attendance systems, save time while ensuring accuracy, though privacy concerns must be considered (see Section 4.2). AI also supports lesson planning by recommending resources aligned with student needs and curriculum goals (Haderer and Ciolacu, 2022). Beyond classroom use, adaptive platforms can personalize professional development for teachers (Yang and Kyun, 2022). Collectively, these productivity gains allow educators to concentrate on meaningful teaching activities and student support.

4.1.3 Enhanced student engagement

Engagement was one of the most frequently reported benefits, as detailed in Table 5. Out of 37 studies, 28 ($\approx 76\%$) indicated that teachers observed higher student motivation and participation when AI tools were integrated into instruction. These included reports of adaptive systems challenging advanced learners while scaffolding struggling students (Yang and Kyun, 2022) and of interactive features sustaining participation through gamification (Chounta et al., 2021). While widespread, these findings remain perception-based and must be interpreted cautiously, as they reflect thematic prevalence across studies rather than outcome-validated effects.

4.2 Concerns and challenges

While teachers frequently acknowledge the potential of AI to enhance learning, their reflections also surface persistent concerns

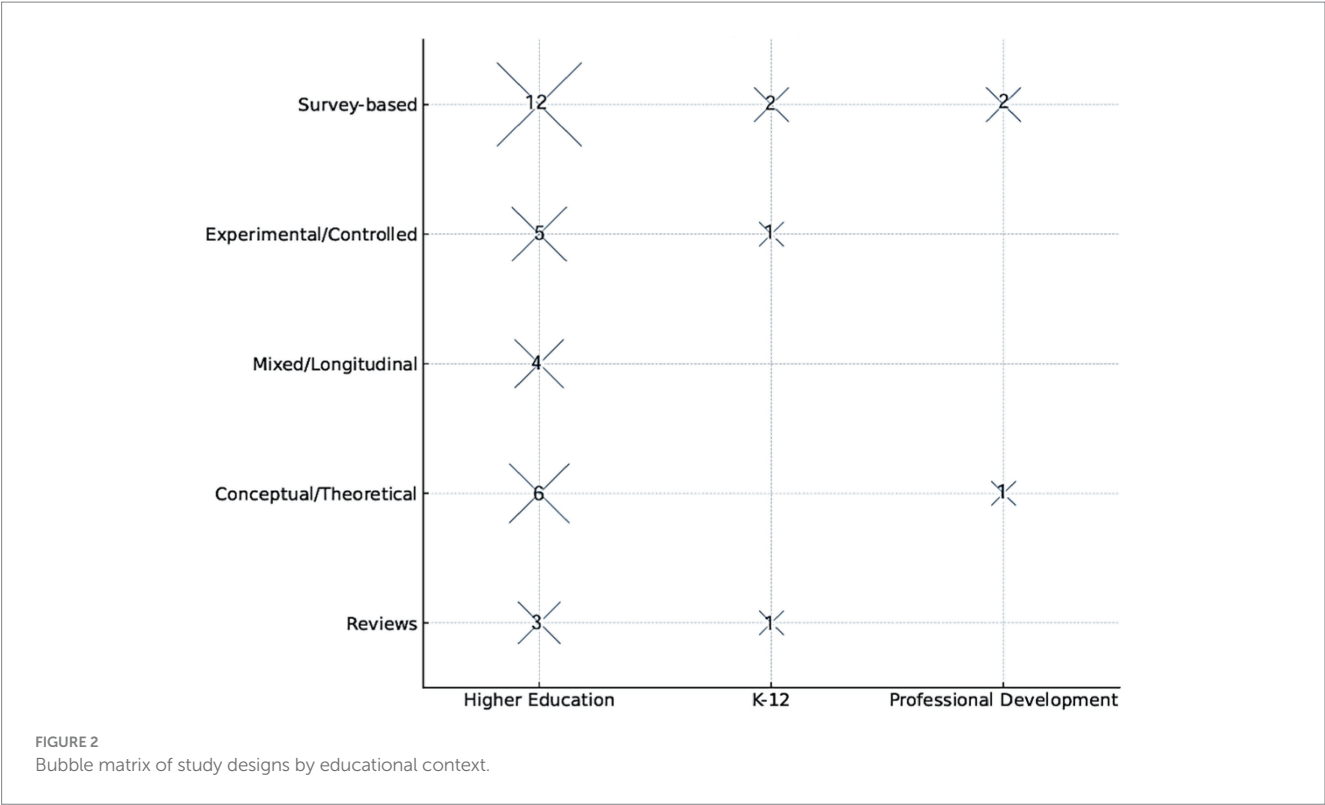


TABLE 4 Distribution of study designs across educational contexts.

Study design	Higher education	K–12	Professional development	Total
Survey-based	12	2	2	16
Experimental/controlled	5	1	0	6
Mixed/longitudinal	4	0	0	4
Conceptual/theoretical	6	0	1	7
Reviews (systematic/meta)	3	1	0	4
Total	30	4	3	37

Bold values indicate the highest frequency within each row, representing the study design most commonly reported in that educational context.

that temper enthusiasm and complicate adoption. Four issues dominate the literature: data privacy and security, algorithmic bias and fairness, relational erosion, and insufficient professional development.

To start with, data privacy and security were among the most frequently cited concerns. Educators highlighted the risks of storing sensitive student information in centralized databases and questioned the adequacy of existing safeguards. These apprehensions were often anticipatory rather than empirically tested, clustering in perception-based studies (Tier D). Yet, they carry practical weight: without strong institutional safeguards, trust in AI systems remains fragile.

Algorithmic bias, on the other hand, emerged as another recurrent theme. Teachers expressed skepticism about the fairness of AI-generated outputs, warning that narrow or non-representative training datasets could entrench existing inequities. In parallel, personalization, the most celebrated benefit, was often described as inseparable from opacity in algorithmic decision-making. The contradiction is stark: the very mechanisms that enable differentiation may also reinforce systemic bias, a tension seldom examined beyond theoretical caution. Relational erosion represents a uniquely pedagogical risk. Although immediacy of feedback is valued for cognitive salience, educators feared that automated responses could substitute for human dialogue, undermining the trust, empathy, and emotional support essential to meaningful teaching. Unlike privacy or bias concerns, which are largely perception-driven, relational erosion was also documented in longitudinal studies (Tier A). These findings suggest that diminished teacher–student interaction is not merely hypothetical but observable over time.

Lack of training and professional development compounds these risks. Teachers consistently noted that without adequate preparation, they risk becoming over-reliant on “black box” systems whose outputs they cannot interpret or challenge. However, this issue is under-investigated in stronger Tier B studies, appearing most prominently in weaker perception-based designs. This imbalance highlights a gap between teachers’ acknowledged needs and the systematic evaluation of interventions that could address them.

Taken together, these concerns demonstrate that benefits and risks are not parallel but interdependent categories. Personalization and immediacy, while celebrated, simultaneously generate new vulnerabilities in fairness and relational trust. Productivity gains can only be sustained if educators are adequately trained to use AI critically, yet evidence of such training remains sparse. Few studies

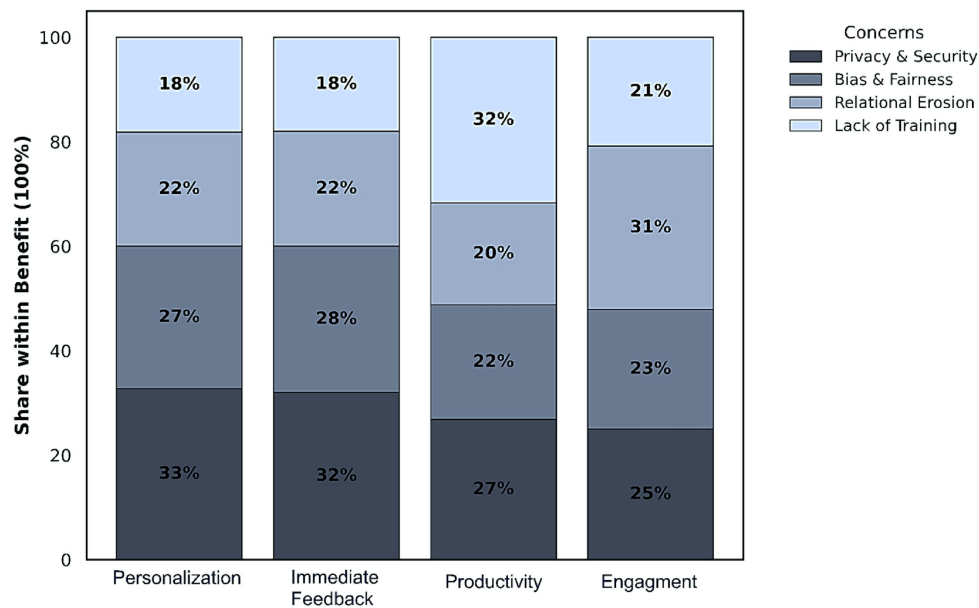


FIGURE 3
Relative frequency of reported benefits vs. concerns in the 37 reviewed studies. Percentages indicate the proportion of studies reporting each theme rather than statistical prevalence across educators.

TABLE 5 Reported benefits of AI-powered feedback in higher education (N = 37 studies).

Benefit	% of studies reporting	No. of studies	Representative evidence
Enhanced personalized instruction	85%	31	AI adapts content to diverse student needs, offering differentiated support.
Immediate and tailored feedback	78%	29	Most teachers report improved learning outcomes from individualized feedback.
Increased productivity	70%	26	Automation of grading, attendance, and scheduling reduces teacher workload.
Enhanced student engagement	76%	28	Teachers observed increased motivation and participation via gamification and tasks.
Support for diverse learners	62%	23	AI adjusts complexity for both advanced and struggling students, enabling inclusivity.

interrogate these contradictions directly, leaving critical questions unresolved: Can AI deliver pedagogical efficiencies without undermining human connection? Can ethical safeguards scale alongside technological innovation? Current evidence provides only partial answers, suggesting that enthusiasm must be tempered by rigorous, outcome-validated inquiry.

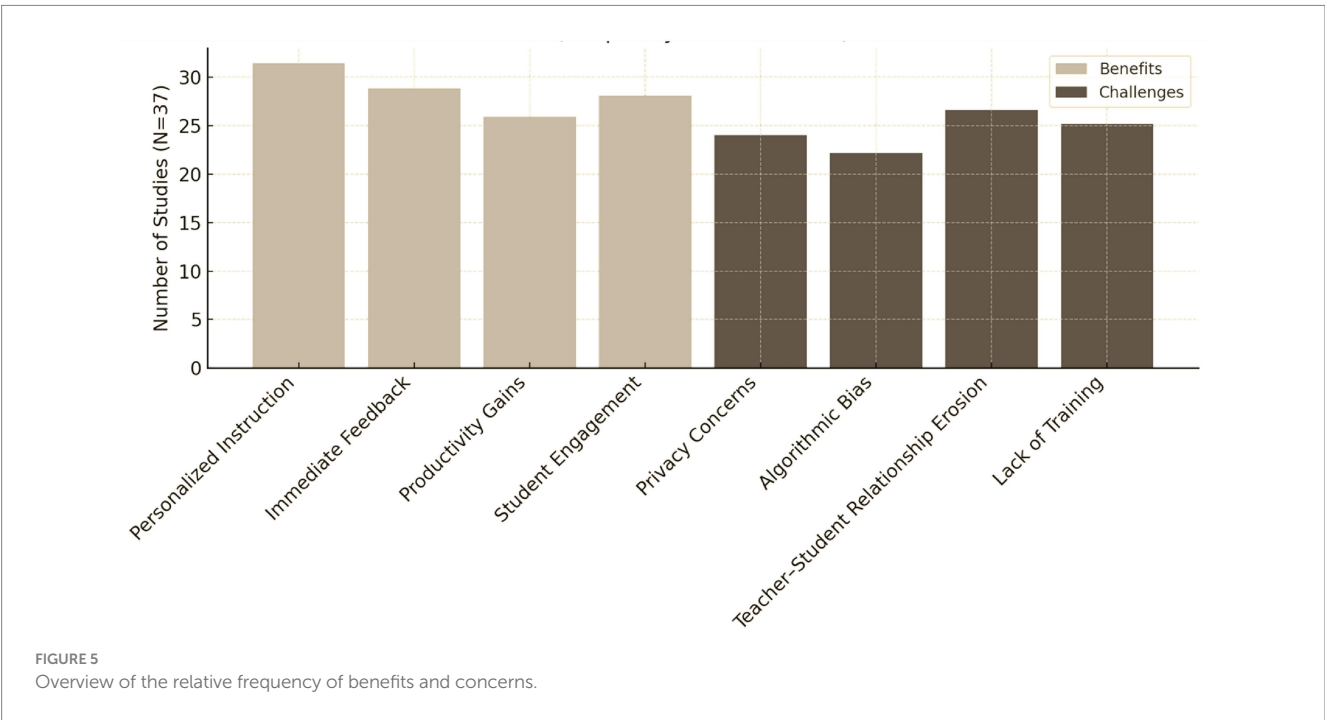
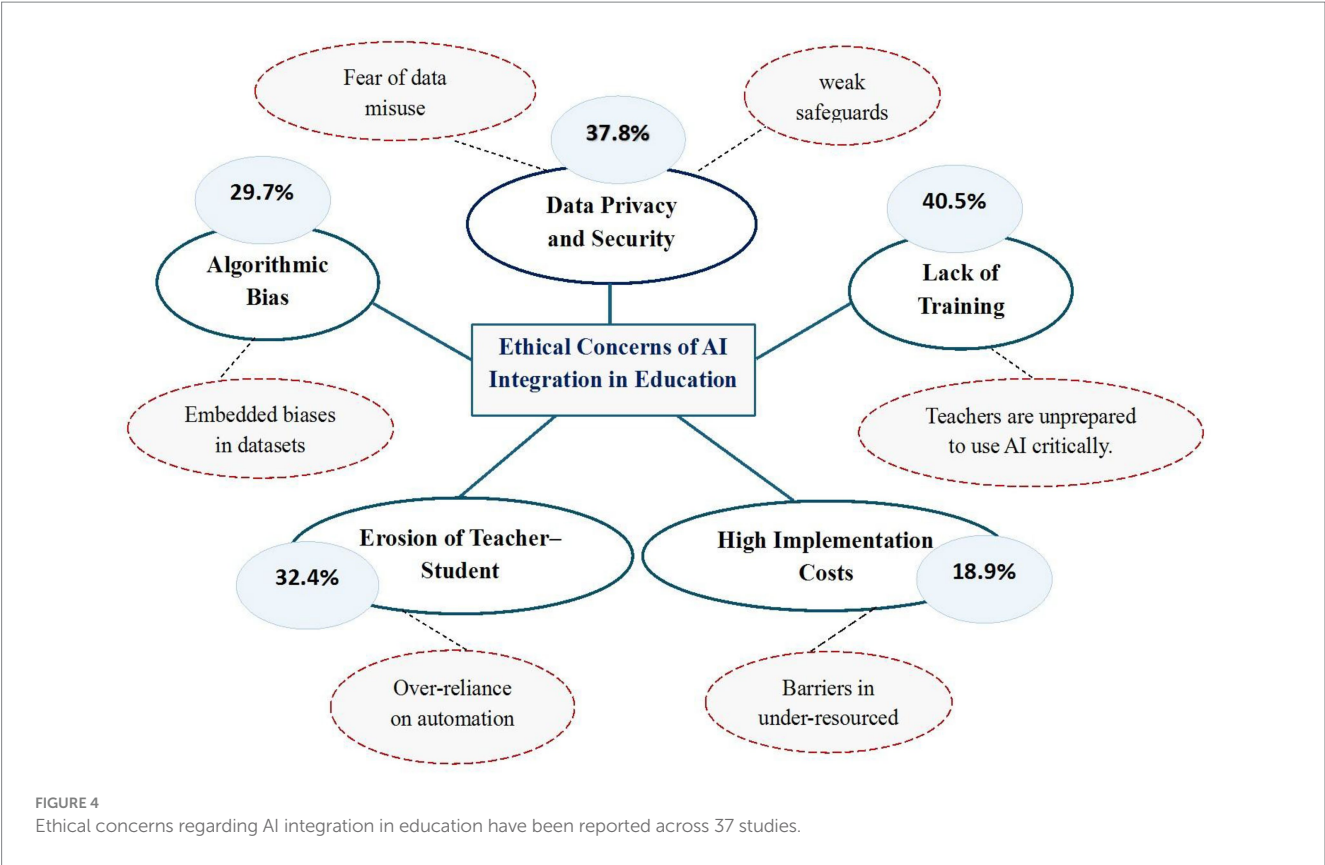
Figure 4 illustrates the relative prevalence of key challenges across the 37 studies. Percentages indicate the proportion of studies identifying each theme, not generalizable statistics about educators. The examples attached to each concern (e.g., weak safeguards,

embedded biases, over-reliance on automation) highlight the intersections between pedagogical risks and ethical vulnerabilities.

The following Figure 5 presents a balanced overview of the positive and negative aspects associated with AI-powered feedback systems in education. The coding analysis revealed that the majority of studies reported positive outcomes, including enhanced personalization, immediacy of feedback, productivity gains, and increased student engagement. At the same time, substantial numbers of studies also highlighted concerns regarding data privacy, algorithmic bias, relational erosion, and insufficient teacher training. These percentages represent the proportion of studies reporting a theme, not generalizable statistics about teachers themselves.

To maintain analytical clarity, findings derived from conceptual or theoretical contributions are reported separately from empirical evidence. Theoretical insights are cited only where they contextualize or interpret empirical results, ensuring that evidence-based outcomes are not conflated with interpretive arguments. In direct answer to RQ4 (obstacles and circumstances affecting successful implementation), the review identifies four recurrent conditions: (i) teacher preparedness and ongoing professional development; (ii) institutional support and governance (including data protection, clear accountability, and transparent tooling); (iii) contextual fit with pedagogy and workload realities; and (iv) infrastructure and resource constraints. Where these conditions are satisfied, educators report sustained uptake and more productive integration; where they are absent, adoption remains tentative or superficial despite short-term enthusiasm.

In Figure 6, the radar chart provides a comparative overview of the relative prevalence of benefits and challenges identified across the 37 included studies. The visualization highlights that *student engagement* and *relational erosion* were among the most frequently reported themes, whereas *high costs* and *productivity gains* were less emphasized. This pattern illustrates the dual role of AI integration in



education. While enhancing personalization, feedback, and learner motivation, it simultaneously raises pressing challenges such as privacy concerns, algorithmic bias, and insufficient teacher training. The balanced representation of benefits and challenges underscores the need for future research to not only capitalize on the pedagogical gains but also to address the systemic and ethical risks associated with the large-scale adoption of AI.

Analysis of the relative prevalence of benefits and challenges across the 37 included studies revealed notable disparities in reporting patterns. The most frequently identified benefits were student

engagement (27 mentions, 73.0%) and *relational erosion* as a recurring challenge (26 mentions, 70.3%). Other highly prevalent themes included *lack of training* (25 mentions, 67.6%), *privacy concerns* (24 mentions, 64.9%), and *personalization* (23 mentions, 62.2%). Moderate representation was observed for *immediate feedback* (22 mentions, 59.5%) and *algorithmic bias* (22 mentions, 59.5%). By contrast, *productivity gains* (20 mentions, 54.1%) and *high costs* (15 mentions, 40.5%) were less frequently emphasized. This distribution suggests that while pedagogical affordances, such as engagement and personalization, are consistently highlighted, ethical and structural risks, including privacy concerns, bias, and relational erosion, are equally pervasive. The relatively lower emphasis on cost-related concerns indicates that financial barriers, though relevant, were not as central to the discourse as the pedagogical and ethical dimensions of AI integration.

4.3 Trends in research activity

The distribution of studies over time highlights a rapid acceleration of interest in AI-automated feedback in recent years. Between 2014 and 2017, empirical studies on this topic were sparse, averaging only one to two publications per year. From 2018 onwards, however, the number of publications increased sharply, coinciding with the broader adoption of adaptive learning platforms and the integration of generative AI systems into educational practice. Figure 7 illustrates the trajectory of relevant publications from 2014 to 2025. The steep upward trend after 2018 reflects the growing recognition of AI as a disruptive force in higher education. This rise also signals both the novelty of the field and the ongoing need for more longitudinal,

large-scale, and cross-contextual studies to consolidate evidence of AI's effectiveness.

In summary, the results of this review paint a dual narrative. On the one hand, AI tools demonstrably support learning personalization, engagement, and efficiency in higher education and online environments. On the other hand, persistent ethical, pedagogical, and trust-related issues accompany these technological advances, indicating that human oversight and strong governance frameworks are required to mediate AI's role in education.

The regional distribution and methodological characteristics of the included studies provide additional context for interpreting these results.

4.4 Regional and cross-cultural distribution

As illustrated in Figure 8, Western and East Asian regions collectively account for nearly three-quarters of the reviewed studies, whereas contributions from the Middle East, Africa, and Latin America remain comparatively limited. The accompanying annotations summarize dominant research orientations ranging from ethical autonomy and accountability in Western contexts to data-driven adaptive-feedback frameworks in East Asian settings. North America contributed the largest share, with 17 studies (45.9%), most situated in higher-education contexts that emphasize teacher autonomy, ethics, and digital feedback control. Europe produced six studies (16.2%), primarily in higher education and teacher-training domains, reflecting a focus on professional development and ethical literacy. East Asia accounted for 10 studies (27.0%), highlighting

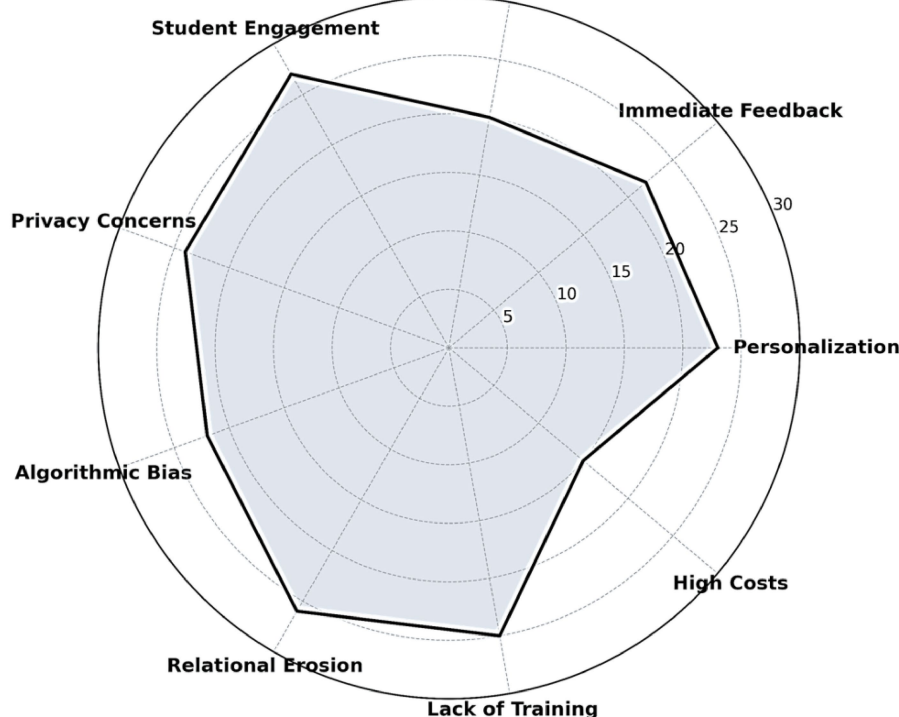


FIGURE 6

Radar chart comparing reported benefits and concerns of AI integration in higher education.

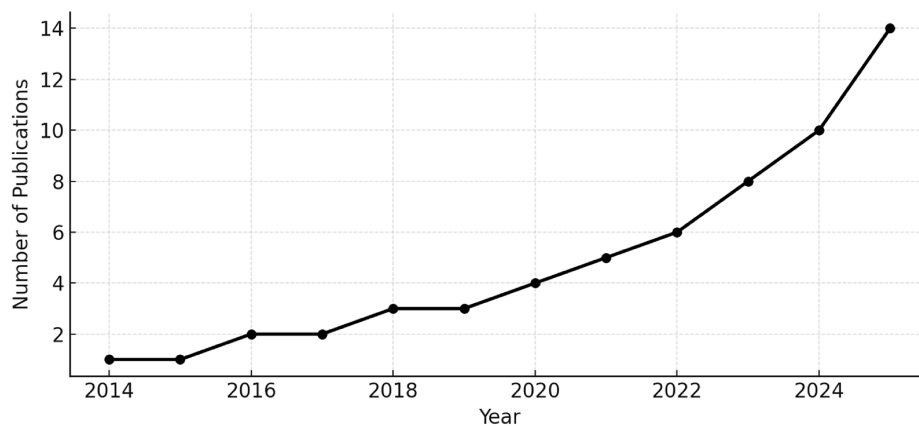


FIGURE 7

The trajectory of relevant publications from 2014 to 2025.

adaptive feedback models, learning analytics, and the pedagogical regulation of AI systems in universities.

The MENA region contributed two studies (5.4%), focusing on AI feedback adoption and cultural adaptation in higher education and STEM-related disciplines. Other regions, including Sub-Saharan Africa, Latin America, and Oceania, together yielded two studies (5.4%), addressing infrastructural readiness and inclusivity in educational technology. Overall, higher education emerges as the dominant research context across regions, yet the clear imbalance in representation underscores the need for more geographically inclusive investigations into AI-mediated feedback practices, particularly within underrepresented Global-South contexts.

5 Discussion

The findings of this review reveal a complex picture of educators' perspectives on AI-automated feedback. Teachers broadly acknowledge the transformative potential of AI to enhance personalized learning, immediacy of feedback, productivity, and student engagement. However, these benefits are consistently accompanied by cautionary concerns about data privacy, ethical risks, algorithmic bias, and diminished teacher–student relationships. This dual stance situates teachers as cautious optimists, simultaneously embracing the promise of AI while foregrounding the conditions required for its responsible use.

Practically, according to research question four, these findings mean that adoption is conditional; implementation succeeds when teacher training, institutional safeguards, pedagogical alignment, and adequate infrastructure co-exist, and falters when any of these are missing.

5.1 Interpreting teachers' perceptions through adoption theories

Teachers' perspectives reveal a dual stance: while efficiency and personalization align with perceived usefulness in adoption theory, ethical reservations, lack of training, and contextual disparities temper optimism. Acceptance is thus conditional, dependent on institutional

support and safeguards. While UTAUT suggests that professional development and institutional support can facilitate adoption, evidence from this review underscores that such acceptance is conditional. Without sufficient training, ethical safeguards, and contextual adaptation, AI systems may increase cognitive load rather than reduce it. Thus, teacher acceptance of AI is better understood as *cautious and context-dependent optimism*; teachers are willing to adopt AI when it demonstrably complements their pedagogy, but this cannot yet be assumed as universal or sustainable.

5.2 Ethical considerations: the FATE framework

The ethical dimensions of AI feedback are aptly captured by the Fairness, Accountability, Transparency, and Ethics (FATE) framework (Memarian and Doleck, 2023). Teachers' concerns about bias and fairness highlight the risk of reproducing educational inequities when algorithms are trained on narrow or unrepresentative datasets. Calls for accountability emphasize that while AI can assist, teachers remain responsible for pedagogical decisions necessitating professional development so that educators can critically interpret and, if necessary, override AI outputs. Transparency is also pivotal: teachers in multiple studies questioned the opacity of algorithmic decision-making and demanded clear disclosure of how AI-generated feedback is produced. Finally, ethics encompasses privacy concerns, as the collection of granular student data by AI systems raises issues of consent and compliance with regulations such as GDPR.

The synthesis suggests that educators' ethical reservations are not abstract; they are practical and influence adoption. Teachers will only integrate AI into daily practice to the extent that these tools align with their ethical standards and the trust they need to have in the technology. In short, teachers want AI systems that are fair in their recommendations, transparent in their functioning, and that clearly leave final decisions in human hands.

5.3 Quality and gaps in the evidence base

Although scholarship on AI-automated feedback in higher education is expanding, the current evidence base remains

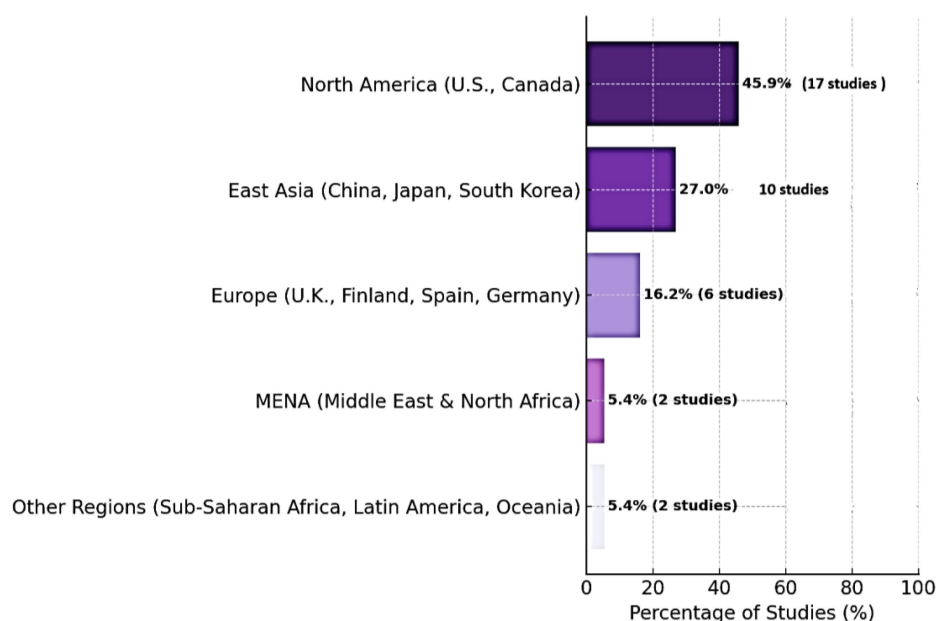


FIGURE 8
Regional distribution by context: summarizing the 37 studies.

methodologically fragile. A large proportion of the reviewed studies rely on self-reported teacher perceptions or short-term pilot interventions. While these approaches provide valuable exploratory insights, they are prone to bias and offer limited evidence of sustained impact. Surveys often capture teachers' attitudes and expectations rather than observable changes in classroom practice, and short-lived experimental projects cannot demonstrate whether AI integration results in lasting pedagogical or institutional transformation.

A critical distinction emerges when comparing perception-based studies (Tier C/D) with outcome-validated designs (Tier A/B). The majority of positive findings, particularly endorsements of personalization, immediacy, productivity, and engagement, stem from surveys or small-scale interventions that capture attitudes and short-term experiences. By contrast, the more rigorous Tier A/B studies, although fewer in number, provide a tempered view: while they confirm efficiency and timeliness benefits, they often report mixed or limited evidence of sustained improvements in achievement or workload reduction. Similarly, ethical concerns such as privacy, bias, and relational erosion are most frequently identified in perception-based studies, raising the possibility that these risks are anticipated rather than empirically observed. This divergence indicates that enthusiasm is largely perception-driven, whereas validated designs yield more cautious or ambivalent results. Explicitly weighing these discrepancies highlights the need for a more critical, evidence-stratified approach to synthesizing findings on AI feedback.

This discrepancy reveals that much of the reported optimism may reflect what can be described as *exploratory enthusiasm* rather than validated transformation. Indeed, contradictions are evident across studies: for instance, teachers who valued AI's immediacy often simultaneously expressed concern about relational erosion or mistrust of algorithmic opacity. Few studies interrogated these tensions systematically, leaving open questions about whether gains in productivity or personalization might come at the cost of diminished teacher-student interaction.

Regional imbalance compounds these limitations. More than half of the reviewed studies were conducted in North America and Europe, while perspectives from the Global South remain underrepresented. This skew limits the generalizability of findings, as cultural traditions, infrastructural realities, and pedagogical norms strongly shape how AI is perceived and used. What appears as a benefit in digitally advanced Western contexts may not replicate in under-resourced or culturally distinct settings.

Taken together, these methodological and regional weaknesses constrain the strength of current evidence. While teachers' perceptions of AI are broadly positive, they cannot yet be taken as proof of enduring improvements in educational outcomes. Establishing such claims will require longitudinal, comparative, and cross-cultural designs that move the field from documenting enthusiasm to validating genuine pedagogical change.

5.4 Teachers as mediators of ethical and pedagogical integration

Perhaps the most consistent theme across the literature is that teachers are not passive recipients of AI systems but active mediators whose acceptance or resistance ultimately shapes AI's effectiveness in practice. Teachers' concerns about professional identity, role redefinition, and job security suggest that AI adoption is deeply intertwined with educators' values and institutional culture, not just technological capability or efficiency.

This review, therefore, reframes AI feedback not as a replacement for human pedagogy but as a relational technology whose impact depends on alignment with educators' practices, ethics, and responsibilities. The evidence suggests that when adequately supported and involved in the process, teachers are willing to integrate AI as a complementary tool. However, without proper safeguards and support, teachers tend to resist the adoption or use of AI in only

superficial ways. In essence, the teacher's role becomes one of guiding the AI, ensuring that it functions in the service of pedagogy and equity, rather than letting the technology dictate the terms of education.

5.5 Comparative weighting of evidence (reframed)

The synthesis of the 37 studies reveals a sharp imbalance between enthusiasm-based evidence and outcome-validated findings. A majority of positive claims regarding personalization, immediacy, and engagement stem from Tier C small-scale interventions and Tier D perception-based surveys (70% of included studies). These studies offer valuable exploratory insights into teachers' optimism and the perceived pedagogical affordances of AI feedback; however, they remain methodologically fragile. By design, they capture short-term attitudes or pilot effects rather than sustained outcomes, and thus risk reflecting what can be termed *exploratory enthusiasm*.

By contrast, only a minority of studies fall into Tier A (longitudinal, large-scale empirical) or Tier B (controlled comparative) designs (30% combined). These more rigorous investigations offer outcome-validated evidence, where teacher perceptions are triangulated with student performance data, comparative benchmarks, or multi-semester follow-ups. Importantly, findings from these tiers are more cautious. While they confirm gains in efficiency and immediacy, they provide weaker or mixed evidence regarding sustained improvements in student achievement or long-term reductions in workload. This suggests that the strongest empirical evidence tempers, rather than amplifies, the optimism prevalent in perception-based research.

To further strengthen claims of methodological fragility, diagnostic reporting was applied where available. Studies classified as Tier A or B typically reported statistical robustness checks, including measures such as adjusted R^2 , AIC/BIC for model fit, and, in some cases, effect sizes. In contrast, Tier C/D studies often lacked such diagnostics, limiting confidence in their generalizability. This discrepancy underscores that the strongest evidence is both quantitatively modest and diagnostically transparent, while the most optimistic claims stem from weaker designs without equivalent rigor.

Taken together, the evidence base reflects not only an imbalance of methodological strength but also a set of unresolved contradictions. Studies that highlight personalization often acknowledge algorithmic opacity in the same breath; reports of efficiency gains frequently coincide with concerns over diminished teacher–student interaction. However, these tensions are rarely interrogated systematically, leaving open whether such trade-offs are temporary features of early adoption or structural consequences of AI integration. Importantly, the stronger Tier A/B studies suggest that while some efficiencies are real, their pedagogical value is contingent and may come at relational or ethical costs. Thus, the synthesis underscores the importance of distinguishing between perception-driven enthusiasm and outcome-validated caution, moving beyond descriptive parallelism to critically evaluate how benefits and risks interact in practice.

6 Conclusion

This systematic review synthesized findings from 37 empirical studies (2014–2024) to explore teachers' perspectives on AI-automated feedback in higher education. The analysis reveals a consistent duality in educators' views: teachers acknowledge AI's capacity to enhance personalized learning, deliver immediate feedback, increase productivity, and foster student engagement, even as they emphasize the parallel risks related to data privacy, algorithmic bias, ethical accountability, and the diminished relational aspects of teaching.

Overall, the evidence suggests that AI-driven feedback is not a neutral tool but rather a pedagogical intervention whose success is shaped by teacher acceptance, ethical safeguards, and institutional support. Teachers' perspectives must therefore be placed at the center of AI integration strategies to ensure that technology adoption reinforces rather than undermines the humanistic values of education.

6.1 Practical and theoretical implications

Continuous professional development is essential for equipping teachers with the skills to interpret and apply AI-generated feedback critically. Rather than replacing pedagogical judgment, AI tools should serve as decision-support systems that complement teachers' expertise. Teachers can use AI to enhance personalization and efficiency, but they must remain the ultimate arbiters of feedback and adapt its use to their students' needs.

The ethical governance frameworks grounded in FATE principles (fairness, accountability, transparency, ethics) are essential for regulating data use, ensuring algorithmic transparency, and establishing clear accountability structures. Guidelines and compliance mechanisms (e.g., alignment with GDPR and student data privacy laws) will help foster teacher trust and willingness to adopt AI tools. Policies should also fund and mandate teacher training programs focused on AI literacy.

Future studies should move beyond exploratory surveys toward longitudinal, comparative, and cross-contextual research designs. It is important to investigate not only how AI feedback affects student learning outcomes but also how it impacts teacher roles, professional identity, and institutional culture. Such research will provide richer insights into the systemic effects of AI in education and help differentiate hype from real, sustained impact.

From a theoretical standpoint, this review contributes by situating teacher perspectives within well-established models of adoption and ethical frameworks, thereby bridging educational technology research with broader theories of innovation, trust, and ethics (In our analysis, concepts from TAM, UTAUT, and FATE were used to anchor the findings, illustrating how teachers' perceived usefulness, trust, and ethical considerations intersect to influence AI adoption).

To advance the field beyond its current exploratory stage, future research must prioritize designs that move from short-term perception-based evidence toward outcome-validated findings. Specifically, three directions stand out as urgent:

- 1 Longitudinal studies with larger samples – Sustained, multi-semester investigations are needed to test whether the reported benefits of AI feedback (e.g., personalization, engagement, and efficiency) endure over time and across diverse cohorts.

- 2 Cross-cultural research in the Global South – Current findings are heavily skewed toward Western higher education. Expanding the evidence base to include underrepresented regions will capture cultural, infrastructural, and pedagogical differences that strongly influence AI adoption and effectiveness.
- 3 Controlled comparative designs – More robust studies are required that directly compare AI-supported feedback with traditional, teacher-led approaches in equivalent contexts. Such comparisons are essential to determine whether improvements in engagement and efficiency are uniquely attributable to AI or whether they could also emerge from conventional pedagogy.

Together, these forward-looking priorities will help transform the evidence base from exploratory enthusiasm into validated knowledge, enabling more confident and context-sensitive decisions about AI integration in higher education.

6.2 Limitations

Despite its contributions, this review faces several limitations. First, the evidence base is methodologically fragile: most studies relied on short-term, self-reported data, with only a small subset providing longitudinal or controlled comparative designs. Second, although inter-rater reliability was high, coding interpretations remain subject to contextual framing, and future reviews should incorporate multi-lingual coders to capture non-English evidence. Third, the literature is regionally imbalanced, with over half of the studies drawn from North America and Europe. This skew limits the transferability of findings to underrepresented regions such as Africa, Latin America, and parts of the Middle East, where cultural traditions, teacher-student relationships, and infrastructural realities may alter both the feasibility and perception of AI tools. In collectivist educational contexts, for example, the relational disruption caused by AI may be more pronounced than in individualist systems. Without cross-cultural validation, the generalizability of current findings remains constrained. In addition, the restriction to English-only publications risks reinforcing Western-centric perspectives and omits potentially important evidence published in languages other than English. Additionally, the inclusion of conceptual and theoretical works alongside empirical studies may dilute the strength of evidence if not carefully distinguished. In this review, theoretical contributions were used to frame and interpret teacher perspectives rather than to provide direct evidence. Nonetheless, separating empirical and non-empirical analyses more explicitly remains an area for refinement in future reviews.

This imbalance risks overstating AI's effectiveness, since much of the current evidence reflects expectations of improvement rather than sustained outcome validation. Additionally, the regional concentration of studies in Western higher education institutions limits the cultural generalizability of the findings. Perspectives from the Global South remain underrepresented, yet are crucial for understanding how diverse educational traditions and infrastructural realities mediate the adoption of AI.

As indicated earlier, caution is warranted when interpreting these findings. The evidence base is heavily skewed toward Western and technologically advanced contexts, and many studies rely on short-term, perception-driven methodologies. These limitations mean that

while teachers frequently report enthusiasm, claims of long-term or generalizable impact remain provisional. Future studies must employ longitudinal, comparative, and cross-cultural designs to test whether the benefits observed in early-adopting contexts extend across diverse educational systems and cultural traditions.

6.3 Key contributions

Unlike earlier systematic reviews that emphasize student learning outcomes or technical performance, this study offers a novel contribution by foregrounding teachers' voices and explicitly bridging adoption models with ethical frameworks, thereby filling a critical gap in the literature. This review makes three distinct contributions to the literature on AI in education:

- (a) Conceptual contribution: It repositions teachers not as passive adopters, but as active mediators of ethical and pedagogical AI integration. By interpreting teacher perspectives through adoption models (TAM/UTAUT), the review expands these models into the educational domain and emphasizes the importance of educator agency in technology use.
- (b) Practical contribution: It identifies concrete prerequisites for successful AI adoption, chiefly the need for ongoing professional development for teachers and the implementation of FATE-based governance frameworks (ensuring fairness, accountability, transparency, and ethics). These are highlighted as non-negotiable conditions for integrating AI in a manner that teachers find acceptable and beneficial.
- (c) Methodological contribution: It exposes the current evidence base's fragility—namely, the lack of long-term, large-scale, and diverse-context studies—and calls for more robust research designs. By doing so, it provides direction for future research efforts, advocating for studies that move beyond self-reported perceptions to measure AI's impact on teaching and learning objectively.

Data availability statement

The original contributions presented in the study are included in the article/[Supplementary material](#), further inquiries can be directed to the corresponding author.

Author contributions

LA: Formal analysis, Investigation, Methodology, Resources, Validation, Visualization, Writing – original draft. TA: Conceptualization, Data curation, Funding acquisition, Project administration, Software, Supervision, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported

and funded by the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University (IMSIU) (grant number IMSIU-DDRSP2504).

Acknowledgments

We extend our sincere gratitude to the educators, researchers, and institutions whose work and insights formed the backbone of this systematic review. Special thanks go to the faculty members and reviewers whose thoughtful feedback guided the refinement of this manuscript. We are deeply grateful to the Deanship of Scientific Research at Imam Mohammad ibn Saud Islamic University for supporting this endeavor through the Graduate Students Research Support Program. Your commitment to advancing educational innovation made this study possible. Finally, to every teacher navigating the evolving landscape of AI in education, you inspire the questions we ask and the solutions we seek.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Abulibdeh, A., Zaidan, E., and Abulibdeh, R. (2024). Navigating the confluence of artificial intelligence and education for sustainable development in the era of industry 4.0: challenges, opportunities, and ethical dimensions. *J. Cleaner Prod.* 437:140527. doi: 10.1016/j.jclepro.2023.140527
- Afzaal, M., Nouri, J., Zia, A., Papapetrou, P., Fors, U., Wu, Y., et al. (2021). Explainable AI for data-driven feedback and intelligent action recommendations to support students self-regulation. *Front. Artif. Intell.* 4:723447. doi: 10.3389/frai.2021.723447
- Atchley, P., Pannell, H., Wofford, K., Hopkins, M., and Atchley, R. A. (2024). Human and AI collaboration in the higher education environment: opportunities and concerns. *Cogn. Res. Princ. Implic.* 9:20. doi: 10.1186/s41235-024-00547-9
- Baidoo-Anu, D., and Ansah, F. (2023). Professional development for AI in education. *Teach. Educ. J.* 27, 205–218. doi: 10.61969/jai.1337500
- Benneh, M. (2023). Artificial intelligence and ethics: a comprehensive review of bias mitigation, transparency, and accountability in AI systems. doi: 10.13140/RG.2.2.23381.19685/1
- Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., et al. (2024). A meta-systematic review of artificial intelligence in higher education: a call for increased ethics, collaboration, and rigour. *Int. J. Educ. Technol. High. Educ.* 21:4. doi: 10.1186/s41239-023-00436-z
- Chen, L., Chen, P., and Lin, Z. (2020). Artificial intelligence in education: a review. *IEEE Access* 8, 75264–75278. doi: 10.1109/ACCESS.2020.2988510
- Chounta, I., Bardone, E., Raudsep, A., and Pedaste, M. (2021). Exploring teachers' perceptions of artificial intelligence as a tool to support their practice in Estonian K-12 education. *Int. J. Artif. Intell. Educ.* 32, 725–755. doi: 10.1007/s40593-021-00243-5
- Crompton, H., and Burke, D. (2023). Artificial intelligence in higher education: the state of the field. *Int. J. Educ. Technol. High. Educ.* 20:22. doi: 10.1186/s41239-023-00392-8
- Doroudi, S. (2022). The intertwined histories of artificial intelligence and education. *Int. J. Artif. Intell. Educ.* 33, 885–928. doi: 10.1007/s40593-022-00313-2
- Escalante, J., Pack, A., and Barrett, A. (2023). AI-generated feedback on writing: insights into efficacy and ENL student preference. *Int. J. Educ. Technol. High. Educ.* 20:57. doi: 10.1186/s41239-023-00425-2
- Fullan, M., Azorin, C., Harris, A., and Jones, M. (2023). Artificial intelligence and school leadership: challenges, opportunities and implications. *Sch. Leadersh. Manage.* 43, 437–444. doi: 10.1080/13632434.2023.2246856
- Haderer, B., and Ciolacu, M. (2022). Education 4.0: artificial intelligence assisted task- and time planning system. *Procedia Comput. Sci.* 200, 1328–1337. doi: 10.1016/j.procs.2022.01.334
- Hamam, D. (2021). "The new teacher assistant: a review of Chatbots' use in higher education" in Communications in Computer and Information Science. HCl international 2021—Posters. eds. C. Stephanidis, M. Antona and S. Ntoa, vol. 1421 (Cham: Springer), 59–63. doi: 10.1007/978-3-030-78645-8_8
- Hew, K. F., Hu, X., Qiao, C., and Tang, Y. (2020). What predicts student satisfaction with MOOCs: a gradient boosting trees supervised machine learning and sentiment analysis approach. *Comput. Educ.* 145:103724. doi: 10.1016/j.compedu.2019.103724
- Holmes, W., Bialik, M., and Fadel, C. (2019). Artificial intelligence in education: promises and implications for teaching and learning. *Center for Curriculum Redesign*. doi: 10.58863/20.500.12424/4276068
- Hooda, M., Rana, C., Dahiya, O., Rizwan, A., and Hossain, M. (2022). Artificial intelligence for assessment and feedback to enhance student success in higher education. *Math. Probl. Eng.* 2022, 1–19. doi: 10.1155/2022/5215722
- Karaca, A., and Kilcan, B. (2023). The adventure of artificial intelligence Technology in Education: comprehensive scientific mapping analysis. *Participatory Educ. Res.* doi: 10.17275/per.23.64.10.4
- Kim, N., and Kim, M. (2022). Teacher's perceptions of using an artificial intelligence-based educational tool for scientific writing. *Front. Educ.* doi: 10.3389/feduc.2022.755914
- Labadze, L., Grigolia, M., and Machaidze, L. (2023). Role of AI chatbots in education: systematic literature review. *Int. J. Educ. Technol. High. Educ.* 20:56. doi: 10.1186/s41239-023-00426-1
- Liang, J., Wang, L., Luo, J., Yan, Y., and Fan, C. (2023). The relationship between student interaction with generative artificial intelligence and learning achievement: serial mediating roles of self-efficacy and cognitive engagement. *Front. Psychol.* 14. doi: 10.3389/fpsyg.2023.1285392
- `, R., and Holmes, W. (2016). Intelligence unleashed: An argument for AI in Education. Open ideas; London: Pearson Education.
- Ma, W., Adesope, O. O., Nesbit, J. C., and Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: a meta-analysis. *J. Educ. Psychol.* 106, 901–918. doi: 10.1037/a0037123
- Memarian, B., and Doleck, T. (2023). Fairness, accountability, transparency, and ethics (FATE) in artificial intelligence (AI), and higher education: a systematic review. *Comput. Educ. Artif. Intell.* 5:100152. doi: 10.1016/j.caeai.2023.100152
- Mollick, E. R., and Mollick, L. (2022). New modes of learning enabled by AI Chatbots: three methods and assignments. doi: 10.2139/ssrn.4300783
- Roll, I., and Wylie, R. (2016). Evolution and revolution in artificial intelligence in education. *Int. J. Artif. Intell. Educ.* 26, 582–599. doi: 10.1007/s40593-016-0110-3
- Rosak-Szyrocka, J., Żywiolek, J., Nayyar, A., and Naved, M. (2023). The role of sustainability and artificial intelligence in education improvement *1st edn.* Chapman and Hall/CRC. doi: 10.1201/9781003425779

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/feduc.2025.1704820/full#supplementary-material>

- Salas-Pilco, S. Z., Xiao, K., and Oshima, J. (2022). Artificial intelligence and new Technologies in Inclusive Education for minority students: a systematic review. *Sustainability*. 14:13572. doi: 10.3390/su142013572
- Samarescu, N., Bumbac, R., Alin, Z., and Iorgulescu, M. (2024). Artificial intelligence in education: next-gen teacher perspectives. *Amfiteatru Economic XXVI*, 145–161. doi: 10.24818/EA/2024/65/145
- Selwyn, N. (2019). Should robots replace teachers? AI and the future of education: Polity.
- Slimi, Z., and Carballido, B. (2023). Systematic review: AI'S impact on higher education - learning, teaching, and career opportunities. *TEM J.* 12:1627. doi: 10.18421/TEM123-44
- Wu, T., He, S., Liu, J., Sun, S., Liu, K., Han, Q. L., et al. (2023). A brief overview of ChatGPT: the history, status quo, and potential future development. *IEEE/CAA J. Automatica Sinica* 10, 1122–1136. doi: 10.1109/JAS.2023.123618
- Xu, W., and Ouyang, F. (2022). The application of AI technologies in STEM education: a systematic review from 2011 to 2021. *Int. J. STEM Ed.* 9:59. doi: 10.1186/s40594-022-00377-5
- Yang, H., and Kyun, S. (2022). The current research trend of artificial intelligence in language learning: a systematic empirical literature review from an activity theory perspective. *Australas. J. Educ. Technol.* 38, 180–210. doi: 10.14742/ajet.7492
- 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? *Int. J. Educ. Technol. High. Educ.* 16, 1–27. doi: 10.1186/s41239-019-0171-0