Association of Researcher of Skills and Vocational Training, Malaysia

ANP-JSSH

ISSN 2773-482X eISSN 2785-8863
DOI: https://doi.org/10.53797/anp.jssh.v7i1.9.2026

An application of Artificial Intelligence in Teaching and Learning: A Systematic Literature Review of Its Implications for University Student and Academic Achievement

Muhammad, Muhammad Adib[1], Benang, Justin Inggang[1], Hajar, Nazran[1], Noor Azan, Muhammad Safwan[1], Setia, Morgan[1], Umpang, Erickson Engga[1] & Ismail, Mohamad Amiruddin[1*]

[1]Faculty of Technical and Vocational Education, Universiti Pendidikan Sultan Idris, 35900 Tanjong Malim, Perak, MALAYSIA

*Corresponding Author Email: amerjp@ftv.upsi.edu.my

Diterima: 10 Ogos 2026; Disemak: 11 Ogos 2026; Diterima untuk penerbitan: 15 September 2026; Tersedia dalam talian: 18 September 2026

Abstract: Artificial intelligence (AI) is increasingly embedded in higher education through adaptive learning systems, intelligent tutoring systems, generative AI assistants and learning analytics, yet evidence on what this integration does to students' motivation and academic achievement remains fragmented. This systematic literature review synthesises twenty studies published between 2022 and 2025 and retrieved from Scopus, following the PRISMA 2020 statement and the Synthesis without Meta-analysis reporting guideline. Records were screened in four stages, appraised against four quality criteria and analysed by thematic synthesis. The corpus is dominated by favourable findings: eighteen of the twenty studies report positive motivational effects, attributed mainly to personal pacing, immediate feedback and increased learner autonomy. The evidentiary basis for those claims, however, is thinner than the pattern suggests. Only half of the included studies report primary empirical data, only four use an experimental design and only one reports a control group comparison, so the apparent consensus rests substantially on conceptual, bibliometric and review papers. Evidence on academic achievement is weaker still, with several studies inferring achievement gains from motivational proxies rather than measuring them. Recurrent risks include over-reliance on generated output, algorithmic bias, data privacy, uneven digital readiness and reduced social presence. The review concludes that AI integration is associated with improved motivation and shows promise for achievement, but that current evidence cannot establish effectiveness. Longitudinal, controlled and discipline-specific studies, together with Malaysian evidence aligned to national generative-AI guidance, are the clear priorities.

Keywords: Artificial Intelligence, Higher Education, Student Motivation, Academic Achievement, Systematic Literature Review

1. Introduction

Artificial Intelligence (AI) has moved from the margins of educational technology to the centre of teaching and learning practice in higher education. Adaptive learning systems, intelligent tutoring systems, conversational agents, automated feedback engines and learning analytics platforms are now routinely embedded in university courses, and the release of general-purpose generative AI tools has accelerated adoption far beyond the pace at which institutional policy has been able to respond. These systems are promoted on the promise of personalised learning pathways, instructional efficiency and sustained learner engagement. Because universities are committing resources on the strength of that promise, the educational consequences of AI integration warrant systematic examination rather than assumption, particularly in relation to the two outcomes institutions most often invoke when justifying investment: student motivation and academic achievement.

The integration of AI in higher education is situated within an established international policy frame. Sustainable Development Goal 4 commits signatory states to inclusive and equitable quality education and lifelong learning opportunities for all (United Nations General Assembly, 2015). The Beijing Consensus on Artificial Intelligence and Education subsequently positioned AI as an instrument for accelerating progress towards that goal, while cautioning that

*Corresponding author: amerjp@ftv.upsi.edu.my

its deployment must be guided by equity, inclusion and transparency (UNESCO, 2019). More recently, UNESCO has issued specific guidance on generative AI in education and research, addressing age thresholds, data governance and the protection of human agency in learning (Miao & Holmes, 2023). These instruments establish that the question is no longer whether AI will feature in higher education, but under what conditions its use can be defended educationally.

Malaysia has moved in step with that international direction. The Malaysia Education Blueprint 2015-2025 (Higher Education) identified globalised online learning as a strategic shift for the university sector (Ministry of Education Malaysia, 2015). The Malaysian Qualifications Agency issued an advisory note on generative AI in higher education in 2023 (Malaysian Qualifications Agency, 2023), and the Ministry of Higher Education subsequently published a national guideline on the use of generative AI technology in higher education teaching and learning (Kementerian Pendidikan Tinggi Malaysia, 2024). Malaysian institutions therefore operate under an explicit regulatory expectation that AI use will be pedagogically justified and academically honest. That expectation raises a prior question that the present review addresses: what does the international evidence actually establish about the learning outcomes such guidance is intended to protect?

Motivation is the construct most frequently invoked in claims made for educational AI, and it requires definition before it can be synthesised. Self-determination theory distinguishes intrinsic motivation, which arises from the inherent satisfaction of an activity, from extrinsic motivation, which is driven by separable outcomes, and holds that motivational quality depends on the extent to which the learning environment supports autonomy, competence and relatedness (Ryan & Deci, 2000, 2020). This distinction matters for the present review because technologies that increase activity, novelty or reward are not thereby supporting autonomy, and a study that reports greater engagement has not necessarily demonstrated a change in motivational quality.

Recent quantitative syntheses give qualified support to the motivational claim. A systematic review and meta-analysis of generative AI found positive pooled effects on student motivation and engagement across educational settings (Xia et al., 2025), and a separate meta-analysis of AI chatbots reported favourable effects on learning outcomes while identifying substantial moderation by intervention design and duration (Wu & Yu, 2024). A further synthesis focused specifically on engagement reached comparable conclusions (Heung & Chiu, 2025). Across these syntheses the direction of effect is consistent, but the magnitude varies widely, and the moderators repeatedly identified are features of the instructional design rather than of the technology itself.

Evidence on academic achievement follows a similar pattern. A meta-analysis restricted to experimental studies found that Chat GPT use was associated with improved learning performance, with effects conditional on task type and the degree of instructional scaffolding (Deng et al., 2025). A second meta-analysis spanning school and university populations reported positive but heterogeneous effects on learning outcomes, with the higher-education subgroup differing from the school subgroup (Liu et al., 2025). Both syntheses caution that positive average effects conceal considerable variation, and neither supports a general claim that AI improves achievement independently of how it is used.

A more sceptical body of evidence qualifies these findings further. A second-order meta-analysis of technology-enhanced learning in higher education, drawing on multiple meta-analyses, concluded that benefits accrue to the learning activities that technology supports rather than to the technology as such, and that substituting a digital tool for conventional instruction yields little by itself (Sailer et al., 2024). An earlier context-specific meta-analysis reached a convergent conclusion, finding that the effects of digital tools depend heavily on implementation conditions and teacher involvement (Hillmayr et al., 2020). Read together, these syntheses shift the explanatory burden from the artefact to its pedagogical deployment.

Two design-level cautions bear directly on how studies in this field should be read. First, perceived learning and measured learning can diverge sharply: in a randomised comparison, students learned substantially more under active instruction while reporting that they had learned less (Deslauriers et al., 2019). Studies that measure motivation or satisfaction by self-report therefore cannot be treated as evidence of learning. Second, motivational gains from technology-supported designs are known to decay: a longitudinal study of gamification found an initial novelty effect that diminished over several weeks before a familiarisation effect emerged (Rodrigues et al., 2022). Short interventions, which predominate in this literature, are especially vulnerable to both problems.

The risks associated with AI use are also becoming better evidenced. A systematic review of over-reliance on AI dialogue systems found associations with weakened critical thinking, analytical reasoning and independent judgement among students (Zhai et al., 2024), and survey evidence links frequent AI tool use to cognitive offloading and lower critical-thinking scores, though that evidence is cross-sectional and self-reported (Gerlich, 2025). Ethical concerns including algorithmic bias, surveillance and data privacy have been articulated as a sector-wide framework for AI in education (Holmes et al., 2022; Akgun & Greenhow, 2022). Underlying several of these risks is a competency question, since AI literacy is an acquired capability rather than a generational endowment (Ng et al., 2021), and the assumption that students arrive technologically competent has been empirically refuted (Kirschner & De Bruyckere, 2017).

Malaysian evidence on these questions is growing but remains fragmented across populations and instruments. Survey work reports that Malaysian undergraduates regard AI tools as useful academic support (Khairuddin et al., 2024) and associate ChatGPT use with improved comprehension while expressing limited trust in its accuracy (Madiah & Abu Samah, 2025). Acceptance studies have examined generative AI adoption among undergraduates using the unified theory of acceptance and use of technology (Jia et al., 2026) and among instructors in terms of confidence and attitudes towards change (Singh et al., 2025), while institutional readiness for AI in the technical and vocational sector has been assessed across staff and student populations (Aida Amalina & Lee, 2025). Within the technical and vocational literature specifically, an AI chatbot deployed as a diagnostic learning support has been reported to raise student test performance (Abd Hamid et al., 2026), digital learning materials have been linked to interest, engagement and achievement (Mohamed Zukri et al., 2026), AI in academic writing has been reviewed (Guo & Zaini, 2024), and knowledge, readiness and attitude have been modelled as predictors of digital learning adoption among teachers (Joseph Thomas et al., 2023).

Despite this accumulating evidence, a gap remains. Existing syntheses tend to isolate a single outcome, examining either motivation or achievement, or to restrict themselves to a single tool class such as chatbots. Few examine motivation and achievement jointly while also accounting for the competency challenges, ethical risks and methodological gaps that condition both, and fewer still read that combined picture against the policy environment now governing Malaysian universities. A synthesis that holds these elements together is needed if institutional decisions are to rest on what the evidence supports rather than on what the technology promises.

This review therefore examines the influence of AI integration on university students' motivation and academic achievement, and identifies the competency challenges, risks and research gaps associated with AI-based teaching and learning. Three research questions are addressed. First, how does AI integration influence university students' motivation? Second, what are the implications of AI for academic achievement and learning outcomes? Third, what challenges, risks and research gaps exist in the implementation of AI in higher education? The review makes no causal claim. It synthesises what a defined corpus reports, appraises the strength of the designs that produced those reports, and states explicitly where the evidence is insufficient to support the conclusions commonly drawn from it.

2. Methodology

2.1 Review Design and Reporting Standards

This study adopted a systematic literature review design to synthesis research on the implications of AI integration for university students' motivation and academic achievement. The review was conducted and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses 2020 statement (Page et al., 2021), which specifies the items required for transparent and replicable reporting of the identification, screening, eligibility and inclusion stages. Because the included studies are methodologically heterogeneous and report outcomes on incommensurable scales, no statistical pooling was attempted. The synthesis is narrative, and is reported in line with the Synthesis Without Meta-analysis guideline, which governs the presentation of reviews that summarise findings by means other than meta-analysis (Campbell et al., 2020). Procedural decisions followed established guidance on conducting systematic reviews in the social sciences (Xiao & Watson, 2019).

2.2 Identification

A systematic search was conducted in the Scopus database, restricted to journal articles published between 2023 and 2025. The search was executed in the TITLE-ABS-KEY field using terms covering artificial intelligence, learning implications, higher education, university students and motivation, and was limited by document type and publication year. An AI-assisted screening tool, Elicit, was used to support the identification stage only; all inclusion decisions, data extraction, analysis and interpretation were carried out manually by the review team. The full search string and its parameters are reproduced verbatim in Table 1. The initial search yielded 310 records.

Table 1: Search String and Search Parameters used in The Systematic Search

Parameter	Specification
Database	Scopus
Search field	TITLE-ABS-KEY
Search string	(TITLE-ABS-KEY (ai AND learning AND implication AND high AND education) AND PUBYEAR > 2022 AND PUBYEAR < 2026)
Publication years	2023 to 2025
Document type	Journal article
Language	English
Records identified	310

Note: The search string is reproduced exactly as executed. The term high rather than higher, and the singular implication, are likely to have reduced recall; this is addressed in section 4.7.

2.3 Screening

Before screening, 126 records were removed. The remaining 184 records were screened on title and abstract, at which stage 129 records were excluded because they did not address AI implementation in an educational setting, did not address ethical or implementation issues, or were not relevant to student competency development. This left 55 records whose full texts were sought and assessed for eligibility.

2.4 Eligibility

The full texts of the 55 remaining records were assessed against predefined inclusion and exclusion criteria, which are set out in Table 2. Studies were included where they addressed AI integration in education with a focus on higher education, examined ethical issues or implementation challenges of AI use, or examined the impact of AI on student competencies or learning outcomes, and were peer-reviewed and indexed in Scopus. At these stage 35 records were excluded, comprising 20 that did not focus on ethical or implementation issues or were not relevant to the review context, and 15 whose publication year fell outside the specified range. Twenty studies were retained for synthesis.

One deviation from the stated protocol is recorded here rather than left for a reader to detect. Lin et al. (2022) falls marginally outside the 2023 to 2025 publication window. It was retained by team consensus because it is the only meta-analysis in the corpus and provides a pre-generative-AI benchmark against which the more recent studies can be read. Its retention is a departure from the protocol and is treated as such in the limitations.

Table 2: Inclusion and Exclusion Criteria Applied at The Eligibility Stage

Criterion	Inclusion	Exclusion
Focus	AI integration in education, with emphasis on higher education	AI research with no educational application
Scope	Addresses ethical issues or implementation challenges of AI use	Does not address ethical or implementation issues
Outcome	Examines impact of AI on student competencies or learning outcomes	Reports no learning-related outcome
Source type	Peer-reviewed and indexed in Scopus	Non-peer-reviewed, grey literature or preprint
Publication year	2023 to 2025	Outside the specified range
Language	English	Non-English

Note: One study published in 2022 was retained by team consensus as the only meta-analysis in the corpus. This deviation is recorded in section 2.4 and in the limitations.

2.5 Quality Appraisal of the Included Studies

The methodological quality and relevance of the 20 included studies were appraised before synthesis. Each study was evaluated against four criteria: clarity of research objectives (C1), methodological clarity (C2), relevance to AI implementation in higher education (C3), and discussion of challenges or implications (C4). Each criterion was scored as met (1) or not met (0), and studies scoring three or above were considered of acceptable quality. All 20 studies met this threshold, and the full scoring is shows in Table 3.

The four criteria were developed by the review team for this review and were not adopted from a validated appraisal instrument such as the Mixed Methods Appraisal Tool (Hong et al., 2018). The scale has not been tested for inter-rater reliability. Both points are limitations and are stated as such rather than presented as methodological provenance. The appraisal does, however, discriminate usefully within the corpus: the ten studies scoring four are those reporting primary empirical data or formal secondary synthesis, while the ten scoring three are conceptual, review or system-design papers that do not report a data collection procedure.

Table 3: Quality Appraisal Scores for The Twenty Included Studies

No.	Study	C1	C2	C3	C4	Total
1	Dan et al. (2025)	1	1	1	1	4
2	Raja et al. (2024)	1	0	1	1	3
3	Nyamwange (2025)	1	1	1	1	4
4	Ren et al. (2024)	1	1	1	1	4
5	Anghel et al. (2025)	1	1	1	1	4
6	Suhag and Singh (2025)	1	0	1	1	3
7	Negi et al. (2024)	1	1	0	1	3
8	Pandoi (2026)	1	0	1	1	3
9	Li et al. (2025)	1	1	1	1	4
10	Zhen (2025)	1	1	1	1	4
11	Lin et al. (2022)	1	1	1	1	4
12	Campillo-Ferrer et al. (2025)	1	1	1	1	4
13	Upadhyay et al. (2025)	1	0	1	1	3
14	Painuli et al. (2025)	1	1	0	1	3

15	Mezquita et al. (2025)	1	0	1	1	3
16	Borna et al. (2024)	1	1	1	1	4
17	Gupta et al. (2024)	1	0	1	1	3
18	Istanti et al. (2024)	1	1	1	1	4
19	Crawford et al. (2024)	1	0	1	1	3
20	Alam (2023)	1	0	1	1	3

Note: C1 = clarity of research objectives; C2 = methodological clarity; C3 = relevance to AI implementation in higher education; C4 = discussion of challenges or implications. Each criterion scored 1 (met) or 0 (not met); threshold for inclusion was a total of 3. C2 was scored 0 where a study reports no data collection procedure. The instrument was developed by the review team and has not been tested for inter-rater reliability; see section 2.5.

2.6 Data Extraction and Thematic Synthesis

Data were extracted manually using a structured framework covering study characteristics, research design, AI application, participants and context, reported motivational findings, reported achievement findings, limitations acknowledged by the authors, and implications for university learning. The complete extraction matrix for all 20 studies is provided in Appendix A. Findings were then analysed by thematic synthesis, in which descriptive themes are derived inductively from the reported findings of the included studies before analytical themes are developed to address the review questions (Thomas & Harden, 2008). The selection process across the identification, screening, eligibility and inclusion stages is summarised in the next section.

3. Results

3.1 Characteristics of the Included Studies

The 20 included studies are heterogeneous in design, and that heterogeneity is itself a finding. Ten studies report primary empirical data: five quantitative surveys or validation studies, four studies described by their authors as experimental, and one learning analytics study. Two are formal secondary syntheses, comprising one meta-analysis and one bibliometric review. The remaining eight are conceptual papers, narrative or critical reviews, or system-design descriptions that report no data collection procedure. Half the corpus, therefore, cannot supply evidence of effect, and any aggregate statement drawn across all 20 studies conflates empirical findings with theoretical propositions.

Design strength within the empirical subset is also limited. Of the four studies described as experimental, only Istanti et al. (2024) explicitly reports a comparison against a control group. No study in the corpus reports random assignment, and none reports a follow-up measurement beyond the intervention period. Sample sizes, where stated, are modest, and several authors identify small samples or single-institution settings as limitations of their own work. Contexts are geographically dispersed, including studies conducted in Kenya, Spain, China, Indonesia and India, and no study in the corpus was conducted in Malaysia.

3.2 Influence of AI Integration on Student Motivation

Across the corpus, 18 of the 20 studies report positive motivational effects associated with AI integration, with Nyamwange (2025) reporting only a moderate increase and Crawford et al. (2024) reporting mixed effects alongside a risk of loneliness. This count is reported as a description of the corpus rather than as an estimate of effect. Counting studies by the direction of their reported findings provides no information about the magnitude of effects and does not weight studies by size or design quality (McKenzie & Brennan, 2024), a caution that applies with particular force here given that half the corpus is non-empirical. The motivational outcomes reported and the studies supporting each are shows in Table 4.

Four mechanisms recur. Personalised pacing and adaptive sequencing were associated with sustained engagement, immediate and private feedback was associated with reduced anxiety and greater willingness to attempt difficult tasks, learner control over pathways was associated with autonomy and self-regulated study, and gamified or interactive elements were associated with increased participation. Notably, the studies reporting the strongest motivational language are disproportionately those without empirical measurement, while the studies that measured motivation with a validated instrument report more qualified gain. Li et al. (2025), which developed and validated a dedicated AI motivation scale, measured motivation only and made no achievement claim.

The corpus also identifies boundary conditions. Nyamwange (2025) found motivational benefit contingent on infrastructure and prior digital skill, Crawford et al. (2024) found that substituting AI for human interaction carried costs to belonging and retention, and Dan et al. (2025) located motivation effects within a rapidly expanding but methodologically uneven research base. Motivational benefit in this literature is therefore conditional rather than automatic.

Table 4: Motivational Outcomes of AI Integration Reported by The Included Studies

Motivational Outcome	Supporting Studies	Reported Mechanism
Increased engagement and participation	Raja et al. (2024); Negi et al. (2024); Borna et al. (2024); Gupta et al. (2024)	Interactive tasks, gamified activity and analytics-driven prompts increased interaction and reduced passive learning.
Enhanced intrinsic motivation	Li et al. (2025); Pandoi (2026); Alam (2023)	Personalised task recommendation and learner-controlled pathways were linked to self-reported intrinsic motivation.
Greater autonomy and self-regulated learning	Zhen (2025); Ren et al. (2024); Suhag and Singh (2025)	Adaptive pacing and immediate feedback supported goal setting, progress monitoring and independent study.
Reduced anxiety and increased confidence	Istanti et al. (2024); Campillo-Ferrer et al. (2025)	Private, non-judgemental feedback allowed self-correction without exposure to peers.
Motivation conditional on guidance and readiness	Nyamwange (2025); Crawford et al. (2024); Dan et al. (2025)	Gains depended on instructor mediation, digital literacy and sustained human presence.

Note: All supporting studies are drawn from the twenty studies listed in Appendix A. Studies appear in more than one row where they report more than one outcome. Row five records qualifying rather than confirming evidence.

3.3 Implications of AI for Academic Achievement

Evidence on academic achievement is considerably weaker than the corpus's favourable tone suggests. Eight studies report some form of achievement evidence, of which the strongest are Lin et al. (2022), a meta-analysis reporting a significant aggregate effect, and Istanti et al. (2024), which reports improvement against a control group. Ren et al. (2024), Negi et al. (2024), Zhen (2025), Anghel et al. (2025), Borna et al. (2024) and Painuli et al. (2025) report gains in test performance, task completion or predictive accuracy relative to observed grades. The remaining twelve studies either state explicitly that achievement was not measured, infer achievement from motivational proxies, or describe a system without evaluating learning. The reported outcomes and their supporting studies are presented in Table 5.

Two features of this evidence warrant emphasis. First, several studies that report achievement benefits are reporting the performance of a predictive model rather than a gain in student learning. Painuli et al. (2025), Borna et al. (2024) and Mezquita et al. (2025) demonstrate that AI systems can identify learning gaps or forecast performance accurately; this is a capability finding, and it establishes the possibility of timely intervention rather than the fact of improved attainment. Second, Campillo-Ferrer et al. (2025) found achievement effects unclear and identified a dependency risk, and Crawford et al. (2024) argued that achievement may decline where AI displaces social learning. The corpus thus contains internal disagreement that the prevailing summary of it obscures.

Table 5: Academic Achievement Outcomes Reported by The Included Studies

Achievement Outcome	Supporting Studies	Reported Mechanism
Higher test or assessment scores	Lin et al. (2022); Istanti et al. (2024); Zhen (2025)	Adaptive difficulty and targeted practice produced measurable gains, in one case against a control group.
Improved comprehension and task completion	Ren et al. (2024); Anghel et al. (2025); Raja et al. (2024)	Personalised pathways and analytics feedback improved conceptual understanding and completion rates.
Early identification of learning gaps	Painuli et al. (2025); Borna et al. (2024); Mezquita et al. (2025)	Predictive models and click-stream analytics flagged at-risk learners, enabling timely intervention.
Efficiency in assessment and feedback	Gupta et al. (2024); Suhag and Singh (2025); Upadhyay et al. (2025)	Automated and adaptive scoring shortened feedback cycles, reported conceptually rather than measured.
Achievement gain not established	Nyamwange (2025); Campillo-Ferrer et al. (2025); Dan et al. (2025); Pandoi (2026); Alam (2023); Crawford et al. (2024)	Reported no direct achievement measure, inferred achievement from motivational proxies, or identified a risk of decline.

Note: Rows three and four record system capability and process efficiency rather than measured student attainment. Row five is reported so that the balance of evidence is visible; it accounts for the larger share of the corpus.

3.4 Challenges, Risks and Research Gaps

While findings were predominantly favourable, the review identified recurrent challenges to AI adoption in higher education, which are categorised with their supporting evidence in Table 6. Five categories emerged: academic integrity and over-reliance on generated output; algorithmic bias and data privacy arising from performance prediction and click-stream analytics; uneven digital readiness and access across student populations; limited instructor preparedness in AI-integrated pedagogy; and reduced social presence were AI substitutes for human interaction.

Five research gaps follow from the corpus. First, no included study reports a longitudinal design, so the durability of motivational and achievement effects is unknown. Second, discipline-specific adoption is under-examined, leaving differences between technical, science and humanities contexts unresolved. Third, research on AI ethics education and responsible-use frameworks is sparse relative to research on AI capability. Fourth, student well-being, including cognitive load and the effort displaced by AI assistance, receives little attention. Fifth, the corpus contains no large-scale comparative study across institutions or countries, and no study conducted in Malaysia, which constrains the transferability of these findings to the Malaysian higher education sector.

Table 6: Challenges, Risks and Competency Gaps Identified in The Included Studies

Category	Challenge Or Risk	Supporting Studies
Academic integrity and over-reliance	Dependence on generated answers may displace independent reasoning and creativity.	Campillo-Ferrer et al. (2025); Crawford et al. (2024)
Algorithmic bias and data privacy	Performance prediction and click-stream analytics raise consent, surveillance and fairness concerns.	Borna et al. (2024); Lin et al. (2022)
Uneven digital readiness and access	Benefits were contingent on infrastructure and on students' prior digital skill.	Nyamwange (2025); Negi et al. (2024)
Instructor preparedness	Limited training in AI-integrated pedagogy constrains implementation quality.	Suhag and Singh (2025); Pandoi (2026); Upadhyay et al. (2025)
Reduced social presence	Substituting AI for human interaction is associated with loneliness and weaker retention.	Crawford et al. (2024); Alam (2023)

Note: Categories were derived inductively by thematic synthesis. All supporting studies are drawn from the twenty studies listed in Appendix A.

4. Discussion

4.1 Motivation Reflects Instructional Design Rather Than Technology

The synthesis supports a qualified conclusion about motivation. The mechanisms most consistently associated with motivational benefit across the corpus are personalised pacing, immediate feedback and learner control over pathways. Read through self-determination theory, these are supports for competence and autonomy rather than properties of AI as a technology (Ryan & Deci, 2020). This interpretation aligns with the second-order meta-analytic finding that gains in technology-enhanced learning accrue to the activity's technology supports rather than to its deployment as such (Sailer et al., 2024), and with the conclusion that effects are conditional on implementation and teacher involvement (Hillmayr et al., 2020).

The practical implication is that an AI tool introduced without a corresponding instructional design should not be expected to reproduce these results. Where the included studies report the largest motivational gains, an instructor has typically restructured feedback timing, task sequencing or learner choice, and the AI system is the means by which that restructuring was delivered. Attributing the outcome to the tool alone misidentifies the active ingredient and encourages procurement decisions that the evidence does not warrant.

4.2 Achievement Evidence Is Thinner Than the Corpus Suggests

The achievement evidence does not currently support claims of effectiveness. Eight of twenty studies report achievement evidence, only one reports a control group comparison, and none reports random assignment or delayed post-testing. This is a materially weaker base than recent meta-analyses of generative AI provide (Deng et al., 2025; Liu et al., 2025), and the divergence is instructive: syntheses restricted to experimental designs report positive but heterogeneous effects that are conditional on task type and scaffolding, whereas the present corpus, containing few controlled designs, presents a more uniformly favourable picture. Uniformity of finding in a methodologically weak corpus is a reason for caution rather than confidence.

Publication processes compound this. Studies reporting null or negative findings are systematically less likely to be published than those reporting positive results (Sterne et al., 2011), so a corpus in which 18 of 20 studies report motivational benefit may reflect what reaches publication as much as what occurs in classrooms. A review that counts favourable findings without acknowledging this cannot distinguish between the two.

4.3 Perceived Benefit Is Not Measured Learning

A distinction that the primary literature frequently elides is central to interpreting these findings: acceptance and perceived benefit are not measured learning. Most motivational findings in this corpus derive from self-report, and self-report is an unreliable proxy for attainment. In a randomised comparison, students learned significantly more under active instruction while reporting that they had learned less (Deslauriers et al., 2019). A student who finds an AI tutor satisfying may be learning more, learning the same, or learning less, and a satisfaction instrument cannot discriminate between those outcomes.

Duration compounds the problem. Motivational gains from technology-supported designs decline after the first weeks of exposure before stabilising (Rodrigues et al., 2022). None of the included studies reports a follow-up measurement, so it cannot be determined whether the reported gains reflect a durable change in motivational quality or the novelty of an unfamiliar tool. This is the single most consequential unknown in the corpus.

4.4 Dependency, Cognitive Offloading and Academic Integrity

The dependency risk identified by five studies in this corpus now has a firmer evidentiary basis than those studies themselves supply. A systematic review of over-reliance on AI dialogue systems reports associations with weakened critical thinking, analytical reasoning and independent judgement among students (Zhai et al., 2024), and correlational evidence links frequent AI use to cognitive offloading and lower critical-thinking performance, though that evidence is cross-sectional and self-reported and should be read accordingly (Gerlich, 2025). The concern raised by Campillo-Ferrer et al. (2025) and Crawford et al. (2024) is therefore not speculative.

This matters for assessment design. If AI assistance displaces the cognitive effort that assessment is intended to elicit, then improved output is compatible with reduced learning, and the achievement gains reported in parts of this corpus become ambiguous. Distinguishing the two requires assessment formats that make the student's reasoning visible, which is a design response rather than a detection problem.

4.5 Ethics, Equity and Institutional Readiness

The ethical and equity concerns identified in the corpus align with the sector-wide framework for AI in education, which treats algorithmic bias, transparency, data governance and human agency as constitutive rather than peripheral issues (Holmes et al., 2022). The privacy concerns arising from click-stream analytics in Borna et al. (2024) are a concrete instance: systems that predict performance from behavioural traces require an explicit consent basis and a defensible position on what inferences may be acted upon (Akgun & Greenhow, 2022).

The digital readiness finding in Nyamwange (2025) deserves particular weight because it contradicts an assumption common in institutional planning. Students do not arrive as uniformly competent technology users; the notion that generational exposure confers competence has been empirically refuted (Kirschner & De Bruyckere, 2017), and AI literacy is an acquired capability requiring deliberate instruction (Ng et al., 2021). Where AI tools are introduced without that instruction, they are likely to widen rather than narrow existing attainment differences.

4.6 Implications for Malaysian Higher Education

For Malaysian higher education, three implications follow. First, the absence of any Malaysian study from a corpus assembled without geographic restriction indicates a genuine gap in the Scopus-indexed literature on AI, motivation and achievement in Malaysian universities, notwithstanding the national survey and acceptance studies that exist (Khairuddin et al., 2024; Madiah & Abu Samah, 2025; Jia et al., 2026; Singh et al., 2025). Second, institutional readiness is a prerequisite rather than a consequence of adoption, a point supported by readiness evidence in the Malaysian technical and vocational sector (Aida Amalina & Lee, 2025) and by findings on teachers' knowledge, readiness and attitudes as predictors of digital learning uptake (Joseph Thomas et al., 2023).

Third, Malaysian practice now has explicit regulatory anchors. The national guideline on generative AI in higher education teaching and learning (Kementerian Pendidikan Tinggi Malaysia, 2024) and the accreditation regulator's advisory note (Malaysian Qualifications Agency, 2023) provide the frame within which institutional policy should be developed, and the present findings suggest what such policy should prioritise: instructional design rather than tool acquisition, assessment formats resistant to cognitive offloading, and explicit AI literacy provision. Recent work in the Malaysian technical and vocational context, including an AI chatbot used for diagnostic learning support (Abd Hamid et al., 2026), digital learning materials linked to interest, engagement and achievement (Mohamed Zukri et al., 2026), and a review of AI in academic writing (Guo & Zaini, 2024), indicates that the local evidence base is beginning to form and would benefit from controlled extension.

4.7 Limitations of This Review

Several limitations constrain what this review can claim. First, a single database was searched. Restricting identification to Scopus excludes relevant work indexed only in Web of Science, ERIC, Scopus-unindexed regional journals and the grey literature, and it is a probable reason why no Malaysian study entered the corpus. Second, the search string was narrow and used the term high rather than higher, which is likely to have reduced recall; the string is reproduced verbatim in Table 1 rather than corrected retrospectively, so that readers can judge its effect.

Third, the corpus is half non-empirical, and the appraisal instrument was developed by the review team rather than adopted from a validated tool, with no inter-rater reliability testing. Fourth, no statistical pooling was possible, so the synthesis reports direction and mechanism rather than magnitude. Fifth, one included study falls outside the stated publication window, as recorded in section 2.4. Sixth, two included studies are book chapters whose bibliographic records could not be independently verified beyond their publishers' own listings, which limits full-text scrutiny. Seventh, screening and extraction were conducted by the review team without a formal double-screening protocol or a reported agreement statistic, and an AI-assisted tool was used at the identification stage, which may have introduced selection effects that cannot be quantified. Finally, only English-language publications were retrieved, excluding Malay-language scholarship relevant to the Malaysian context.

4.8 Directions for Future Research

Each recommendation below answers a limitation named above. To address the absence of longitudinal evidence, studies should measure motivation and achievement at multiple points beyond the intervention period, so that novelty effects can be distinguished from durable change. To address the scarcity of controlled designs, future work should employ comparison groups and, where ethically feasible, random assignment, reporting effect sizes with confidence intervals rather than statements of significance alone.

To address the conflation of perception with attainment, studies should pair validated motivation instruments with independent achievement measures administered under conditions in which AI assistance is controlled. To address the single-database limitation and the absence of Malaysian evidence, future reviews should search multiple databases including regional and Malay-language sources, and primary studies should be conducted in Malaysian universities under the national generative-AI guideline. To address the dependency risk, research should examine assessment formats that make student reasoning visible, and should measure the cognitive effort displaced by AI assistance rather than output quality alone. Finally, discipline-specific studies are needed to establish whether the mechanisms identified here operate comparably across technical, scientific and humanities programmes.

5. Conclusion

This systematic literature review examined what the recent literature establishes about the implications of artificial intelligence for university students' motivation and academic achievement, and about the challenges, risks and research gaps that accompany its use. The corpus positions AI as a broadly favourable presence in higher education, associated with improved motivation through personalised pacing, immediate feedback and greater learner autonomy, and showing promise for achievement through adaptive sequencing and the early identification of learning gaps.

The strength of that evidence, however, does not match the confidence with which it is commonly reported. Half the corpus reports no primary data, controlled designs are rare, follow-up measurement is absent, and several studies infer achievement from motivational proxies rather than measuring it. The mechanisms that appear to matter are features of instructional design, notably feedback timing, task sequencing and learner control, rather than properties of the technology itself. Recurrent risks including over-reliance on generated output, algorithmic bias, data privacy, uneven digital readiness and reduced social presence qualify the favourable picture further.

The defensible conclusion is that AI integration is associated with improved student motivation and holds promise for academic achievement, but that current evidence cannot establish effectiveness. For Malaysian higher education, where national guidance on generative AI is now in place, the priority is not faster adoption but better evidence: longitudinal and controlled studies, conducted locally, that measure learning as well as perception, and institutional policy that invests in instructional design and AI literacy rather than in tools alone.

Acknowledgement

The authors thank the Faculty of Technical and Vocational, Universiti Pendidikan Sultan Idris, for institutional support during the preparation of this review. This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Conflict of Interest

The authors declare no conflicts of interest.

Appendix A: Data Extraction Matrix

The complete data extraction matrix for the twenty included studies is presented in Table A1. Columns follow the structured extraction framework described in section 2.6.

Table A1: Data Extraction Matrix for The Twenty Included Studies

No.	Study	Design	AI Application	Context	Motivation Finding	Achievement Finding
1	Dan et al. (2025)	Bibliometric	General AI trends	Global dataset	Research base expanding; strong reported link to motivation.	Indirect; inferred through motivation pathways.
2	Raja et al. (2024)	Qualitative review	Intelligent tutoring systems	Higher education	AI reported to raise interest and engagement.	Better work quality and task completion; no empirical measurement.
3	Nyamwange (2025)	Survey	AI-enabled learning management	Universities, Kenya	Moderate motivational increase.	No direct achievement metric reported.
4	Ren et al. (2024)	Experimental	Adaptive learning systems	Students, China	Real-time feedback enhanced motivation.	Achievement improved through personalised pathways; small sample.
5	Anghel et al. (2025)	Mixed methods	AI learning analytics	Students, Romania	Motivation increased via interactive content.	Improved learning outcomes; context-specific.
6	Suhag and Singh (2025)	Conceptual	AI in blended learning	Higher education	Proposed increase in participation and curiosity.	Proposed gain in task success; no field data.
7	Negi et al. (2024)	Experimental	AI learning apps	First-year students, India	High engagement attributed to gamification.	Learning performance improved; focus not purely tertiary.
8	Pandoi (2026)	Narrative review	Creative AI systems	Higher education	Stimulation of intrinsic motivation proposed.	Achievement potential inferred; limited evidence.
9	Li et al. (2025)	Scale validation	AI-enabled learning tasks	300 students	AI enhanced intrinsic and extrinsic motivation.	Not measured; motivation only.
10	Zhen (2025)	Survey, path model	AI English tutoring	Students, China	Strong positive motivational effects.	Autonomy associated with improved test scores; domain-specific.
11	Lin et al. (2022)	Meta-analysis	Multiple AI models	45 datasets	AI consistently associated with higher motivation.	Significant aggregate achievement effect; design variation noted.
12	Campillo-Ferrer et al. (2025)	Quantitative survey	Generative AI tools	Students, Spain	Students reported greater motivation with AI support.	Achievement unclear; over-reliance identified.
13	Upadhyay et al. (2025)	Review, case	AI teaching assistants	Higher education	Motivation reported to rise through easier task handling.	Performance improvement asserted; conceptual.

14	Painuli et al. (2025)	Experimental	Prediction model	Computing students	Motivation through progress tracking.	High prediction accuracy against grades; technical context.
15	Mezquita et al. (2025)	System design	AI academic platform	University	Real-time feedback proposed as motivating.	Performance monitoring capability; scalability unknown.
16	Borna et al. (2024)	Learning analytics	Clickstream data	Online students	Behavioural tracking associated with engagement.	Predicts achievement effectively; privacy concerns.
17	Gupta et al. (2024)	Conceptual framework	Generative AI assessment	Universities	Higher engagement proposed post-pandemic.	Adaptive scoring proposed; no empirical dataset.
18	Istanti et al. (2024)	Experimental	AI adaptive tutoring	Students, Indonesia	Motivation driven by personalisation.	Significant improvement against control group; language-only sample.
19	Crawford et al. (2024)	Critical review	Human replacement	Higher education	Motivation mixed; loneliness risk identified.	Achievement may decline without social learning.
20	Alam (2023)	Analytical review	AI learning ecologies	Higher education	Potential to increase autonomy-related motivation.	Long-term achievement improvement proposed; theoretical.

Note: Designs are reported as described by the original authors. Entries in the motivation and achievement columns summarise what each study reports, not what this review endorses. Studies 6 and 13 are book chapters whose bibliographic records could not be independently verified beyond their publishers' listings.

References

Abd Hamid, M. S., Marjuni, A., Nasruddin, M. A., & Abdullah, M. S. (2026). Penggunaan AI chatbot sebagai alat sokongan pembelajaran diagnostik kerosakan enjin dalam kalangan pelajar teknologi automotif kolej vokasional: Satu kajian tindakan. *ANP Journal of Social Sciences and Humanities, 7*(1), 69-82. https://doi.org/10.53797/anp.jssh.v7i1.7.2026

Aida Amalina, & Lee, M. F. (2025). Kesediaan institusi TVET ke arah penggunaan teknologi kecerdasan buatan dalam era digital. *Online Journal for TVET Practitioners, 10*(2), 34-42. https://journal.uthm.edu.my/index.php/oj-tp/article/view/20809

Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics, 2*(3), 431-440. https://doi.org/10.1007/s43681-021-00096-7

Alam, A. (2023). Intelligence unleashed: An argument for AI-enabled learning ecologies with real world examples of today and a peek into the future. *AIP Conference Proceedings, 2717*(1), Article 030001. https://doi.org/10.1063/5.0129803

Anghel, G. A., Zanfir, C. M., Matei, F. L., Voicu, C. D., & Neacsa, R. A. (2025). The integration of artificial intelligence in academic learning practices: A comprehensive approach. *Education Sciences, 15*(5), Article 616. https://doi.org/10.3390/educsci15050616

Borna, M.-R., Saadat, H., Hojjati, A. T., & Akbari, E. (2024). Analyzing click data with AI: Implications for student performance prediction and learning assessment. *Frontiers in Education, 9*, Article 1421479. https://doi.org/10.3389/feduc.2024.1421479

Campbell, M., McKenzie, J. E., Sowden, A., Katikireddi, S. V., Brennan, S. E., Ellis, S., Hartmann-Boyce, J., Ryan, R., Shepperd, S., Thomas, J., Welch, V., & Thomson, H. (2020). Synthesis without meta-analysis (SWiM) in systematic reviews: Reporting guideline. *BMJ, 368*, l6890. https://doi.org/10.1136/bmj.l6890

Campillo-Ferrer, J. M., Lopez-Garcia, A., & Miralles-Sanchez, P. (2025). Student perceptions of the use of Gen-AI in a higher education program in Spain. *Digital, 5*(3), Article 29. https://doi.org/10.3390/digital5030029

Crawford, J., Allen, K.-A., Pani, B., & Cowling, M. (2024). When artificial intelligence substitutes humans in higher education: The cost of loneliness, student success, and retention. *Studies in Higher Education, 49*(5), 883-897. https://doi.org/10.1080/03075079.2024.2326956

Dan, L., Mohamed, H. B., & Abuhassna, H. (2025). A bibliometric analysis of artificial intelligence: Impact on student motivation in education. In *AI in education, governance, and leadership: Adoption, impact, and ethics* (pp. 97-120). IGI Global. https://doi.org/10.4018/979-8-3373-5550-4.ch005

Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2025). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. *Computers & Education, 227*, Article 105224. https://doi.org/10.1016/j.compedu.2024.105224

Deslauriers, L., McCarty, L. S., Miller, K., Callaghan, K., & Kestin, G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *Proceedings of the National Academy of Sciences, 116*(39), 19251-19257. https://doi.org/10.1073/pnas.1821936116

Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies, 15*(1), Article 6. https://doi.org/10.3390/soc15010006

Guo, H., & Zaini, S. H. (2024). Artificial intelligence in academic writing: A literature review. *Asian Pendidikan, 4*(2), 46-55. https://doi.org/10.53797/aspen.v4i2.6.2024

Gupta, S., Dharamshi, R. R., & Kakde, V. (2024). An impactful and revolutionized educational ecosystem using generative AI to assist and assess the teaching and learning benefits, fostering the post-pandemic requirements. In *2024 2nd International Conference on Emerging Trends in Information Technology and Engineering (IC-ETITE)* (pp. 1-4). IEEE. https://doi.org/10.1109/ic-ETITE58242.2024.10493370

Heung, Y. M. E., & Chiu, T. K. F. (2025). How ChatGPT impacts student engagement from a systematic review and meta-analysis study. *Computers and Education: Artificial Intelligence, 8*, Article 100361. https://doi.org/10.1016/j.caeai.2025.100361

Hillmayr, D., Ziernwald, L., Reinhold, F., Hofer, S. I., & Reiss, K. M. (2020). The potential of digital tools to enhance mathematics and science learning in secondary schools: A context-specific meta-analysis. *Computers & Education, 153*, Article 103897. https://doi.org/10.1016/j.compedu.2020.103897

Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Buckingham Shum, S., Santos, O. C., Rodrigo, M. M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2022). Ethics of AI in education: Towards a community-wide framework. *International Journal of Artificial Intelligence in Education, 32*(3), 504-526. https://doi.org/10.1007/s40593-021-00239-1

Hong, Q. N., Fabregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M.-C., Vedel, I., & Pluye, P. (2018). The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Education for Information, 34*(4), 285-291. https://doi.org/10.3233/EFI-180221

Istanti, W., Suseno, Pratiwi, S., & Saddhono, K. (2024). AI-driven personalized learning: Revolutionizing language education. In *2024 International Conference on IoT, Communication and Automation Technology (ICICAT)* (pp. 329-334). IEEE. https://doi.org/10.1109/ICICAT62666.2024.10923274

Jia, B., Mohamad, M., Kamsin, I. F., & Shim, J. (2026). Generative artificial intelligence-powered learning: A UTAUT study among undergraduate students from a public university in Malaysia. *Arab World English Journal (AWEJ) Special Issue on CALL, 17*(12), 288-304. https://doi.org/10.24093/awej/call12.17

Joseph Thomas, J. S., Jamaluddin, R., & Abdullah, A. (2023). The use of digital learning: A predictive model of knowledge, readiness and attitude factors among Technical and Vocational Education (TVET) teachers in secondary schools. *ANP Journal of Social Sciences and Humanities, 4*(2), 32-40. https://doi.org/10.53797/anp.jssh.v4i2.5.2023

Kementerian Pendidikan Tinggi Malaysia. (2024). *Garis panduan penggunaan teknologi kecerdasan buatan generatif (KBG) dalam pengajaran dan pembelajaran (PdP) pendidikan tinggi*. Jabatan Pendidikan Tinggi, Kementerian Pendidikan Tinggi Malaysia. https://repositori.mohe.gov.my/109/

Khairuddin, Z., Shahabani, N. S., Ahmad, S. N., Ahmad, A. R., & Zamri, N. A. (2024). Students' perceptions on the artificial intelligence (AI) tools as academic support. *Malaysian Journal of Social Sciences and Humanities, 9*(11), Article e003087. https://doi.org/10.47405/mjssh.v9i11.3087

Kirschner, P. A., & De Bruyckere, P. (2017). The myths of the digital native and the multitasker. *Teaching and Teacher Education, 67*, 135-142. https://doi.org/10.1016/j.tate.2017.06.001

Li, J., King, R. B., Chai, C. S., Zhai, X., & Lee, V. W. Y. (2025). The AI Motivation Scale (AIMS): A self-determination theory perspective. *Journal of Research on Technology in Education, 58*(5), 961-982. https://doi.org/10.1080/15391523.2025.2478424

Lin, R., Zhang, Q., Xi, L., & Chu, J. (2022). Exploring the effectiveness and moderators of artificial intelligence in the classroom: A meta-analysis. In J. Yang, D. Liu, Kinshuk, A. Tlili, M. Chang, E. Popescu, D. Burgos, & Z. Altinay (Eds.), *Resilience and future of smart learning: Proceedings of 2022 International Conference on Smart Learning Environments* (pp. 61-66). Springer. https://doi.org/10.1007/978-981-19-5967-7_7

Liu, X., Guo, B., He, W., & Hu, X. (2025). Effects of generative artificial intelligence on K-12 and higher education students' learning outcomes: A meta-analysis. *Journal of Educational Computing Research, 63*(5), 1249-1291. https://doi.org/10.1177/07356331251329185

Madiah, M., & Abu Samah, A. R. (2025). Exploring the relationship between academic achievement and ChatGPT usage: A survey of higher education students in Malaysia. *Journal of Informatics and Web Engineering, 4*(2), 275-302. https://doi.org/10.33093/jiwe.2025.4.2.18

Malaysian Qualifications Agency. (2023). *Advisory note no. 2/2023: Penggunaan teknologi kecerdasan buatan generatif (generative artificial intelligence) dalam pendidikan tinggi*. Agensi Kelayakan Malaysia.

McKenzie, J. E., & Brennan, S. E. (2024). Chapter 12: Synthesizing and presenting findings using other methods. In J. P. T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, & V. A. Welch (Eds.), *Cochrane handbook for systematic reviews of interventions* (Version 6.5). Cochrane. https://training.cochrane.org/handbook/current/chapter-12

Mezquita, Y., Parra, J., Alonso-Rincon, R., & Prieto, J. (2025). Data-driven academic performance evaluation: A smart platform approach. In M. F. Santos, J. Machado, P. Novais, P. Cortez, & P. M. Moreira (Eds.), *Progress in artificial intelligence: 23rd EPIA Conference on Artificial Intelligence, EPIA 2024, proceedings, part I* (Lecture Notes in Computer Science, Vol. 14967, pp. 105-114). Springer. https://doi.org/10.1007/978-3-031-73497-7_9

Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000386693

Ministry of Education Malaysia. (2015). *Malaysia education blueprint 2015-2025 (higher education)*. Ministry of Education Malaysia. https://repositori.mohe.gov.my/id/eprint/105/

Mohamed Zukri, N. I., Mohd Zakee, N. N., Ahmad Hambali, I. A., & Faizal, A. E. (2026). Penggunaan bahan pembelajaran digital dalam sesi pembelajaran teori di kolej vokasional: Satu kajian tindakan terhadap minat, penglibatan dan pencapaian pelajar. *ANP Journal of Social Sciences and Humanities, 7*(1), 47-57. https://doi.org/10.53797/anp.jssh.v7i1.5.2026

Negi, P., Shrivastava, V. K., Pandey, S., Chythanya, K. R., Negi, P., & Gupta, M. (2024). Intervention of artificial intelligence in education to reduce illiteracy rate. In *2024 3rd International Conference on Sentiment Analysis and Deep Learning (ICSADL)* (pp. 153-157). IEEE. https://doi.org/10.1109/ICSADL61749.2024.00031

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence, 2*, Article 100041. https://doi.org/10.1016/j.caeai.2021.100041

Nyamwange, C. (2025). Assessing the extent of integration of artificial intelligence in teaching and learning at Kenyan universities. *Pan-African Journal of Education and Social Sciences, 6*(1), 76-87. https://doi.org/10.56893/pajes2025v06i01.06

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hrobjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71

Painuli, P., Rawat, A. J., Dabral, U., & Devi, H. (2025). AI in education: Developing a student performance prediction engine using ML techniques. In *2025 World Skills Conference on Universal Data Analytics and Sciences (WorldSUAS)* (pp. 1-6). IEEE. https://doi.org/10.1109/WorldSUAS66815.2025.11198942

Pandoi, D. (2026). The role of artificial intelligence in futuristic education models: Transforming learning and empowering students. In S. Dua, S. Dua, & P. Kaur (Eds.), *Transforming business education through artificial intelligence*. Apple Academic Press.

Raja, S., Jebadurai, D. J., Lavrik, I., Rudenko, M., Kalenychenko, R., & Reznik, N. (2024). Impact of artificial intelligence in students' learning life. In R. Khamis & A. Buallay (Eds.), *AI in business: Opportunities and limitations: Volume 2* (Studies in Systems, Decision and Control, Vol. 516, pp. 3-17). Springer. https://doi.org/10.1007/978-3-031-49544-1_1

Ren, Y., Jiang, W., & Lu, H. (2024). Analysis on the application of artificial intelligence technology in college teaching. In Y. Zhang & N. Shah (Eds.), *Application of big data, blockchain, and Internet of Things for education informatization* (Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, Vol. 582, pp. 434-444). Springer. https://doi.org/10.1007/978-3-031-63136-8_45

Rodrigues, L., Pereira, F. D., Toda, A. M., Palomino, P. T., Pessoa, M., Carvalho, L. S. G., Fernandes, D., Oliveira, E. H. T., Cristea, A. I., & Isotani, S. (2022). Gamification suffers from the novelty effect but benefits from the familiarization effect: Findings from a longitudinal study. *International Journal of Educational Technology in Higher Education, 19*(1), Article 13. https://doi.org/10.1186/s41239-021-00314-6

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68-78. https://doi.org/10.1037/0003-066X.55.1.68

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology, 61*, Article 101860. https://doi.org/10.1016/j.cedpsych.2020.101860

Sailer, M., Maier, R., Berger, S., Kastorff, T., & Stegmann, K. (2024). Learning activities in technology-enhanced learning: A systematic review of meta-analyses and second-order meta-analysis in higher education. *Learning and Individual Differences, 112*, Article 102446. https://doi.org/10.1016/j.lindif.2024.102446

Singh, P., Singh, H., Somasundram, P., & Anthonysamy, L. (2025). Exploring the role of instructors' confidence and attitudes towards change in the acceptance of ChatGPT as the future of teaching and learning. *Asian Journal of University Education, 21*(3), 695-709. https://doi.org/10.24191/ajue.v21i3.47

Sterne, J. A. C., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., Carpenter, J., Rucker, G., Harbord, R. M., Schmid, C. H., Tetzlaff, J., Deeks, J. J., Peters, J., Macaskill, P., Schwarzer, G., Duval, S., Altman, D. G., Moher, D., & Higgins, J. P. T. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ, 343*, d4002. https://doi.org/10.1136/bmj.d4002

Suhag, N., & Singh, A. (2025). Blended learning in the age of artificial intelligence: Transforming education through innovation. In *Reshaping blended learning environments with AI*. IGI Global.

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology, 8*(1), Article 45. https://doi.org/10.1186/1471-2288-8-45

UNESCO. (2019). *Beijing consensus on artificial intelligence and education*. https://unesdoc.unesco.org/ark:/48223/pf0000368303

United Nations General Assembly. (2015). *Transforming our world: The 2030 agenda for sustainable development* (A/RES/70/1). United Nations. https://sdgs.un.org/2030agenda

Upadhyay, P., Bhardwaj, H., & Tomar, P. (2025). Artificial intelligence tools for instructors and learners to optimize the teaching and learning processes. In *AI based solutions for inclusive quality education*. IGI Global.

Wu, R., & Yu, Z. (2024). Do AI chatbots improve students' learning outcomes? Evidence from a meta-analysis. *British Journal of Educational Technology, 55*(1), 10-33. https://doi.org/10.1111/bjet.13334

Xia, Q., Li, W., Yang, Y., Weng, X., & Chiu, T. K. F. (2025). A systematic review and meta-analysis of the effectiveness of Generative Artificial Intelligence (GenAI) on students' motivation and engagement. *Computers and Education: Artificial Intelligence, 9*, Article 100455. https://doi.org/10.1016/j.caeai.2025.100455

Xiao, Y., & Watson, M. (2019). Guidance on conducting a systematic literature review. *Journal of Planning Education and Research, 39*(1), 93-112. https://doi.org/10.1177/0739456X17723971

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. *Smart Learning Environments, 11*(1), Article 28. https://doi.org/10.1186/s40561-024-00316-7

Zhen, X. (2025). Application of artificial intelligence in English learning: The interactive effects of teacher support, learning motivation, and autonomous learning ability. *Acta Psychologica*, *261*, Article 105960. https://doi.org/10.1016/j.actpsy.2025.105960