INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN PROGRESSIVE EDUCATION & DEVELOPMENT

OPEN ACCESS JOURNAL

Vol 15, Issue 4, (2026) E-ISSN: 2226-6348

Mapping the Evidence on AI Dependence, Self-Efficacy, and Academic Achievement in Higher Education: A Scoping Review (2015–2025)

Arafah Omer Mohammed Ahmed[1], Asad Ur Rehman Muhammad Sadiq[2]

[1]Graduate School of Management, Management and Science University (MSU), Malaysia,
[2]Faculty of Business and Management, Management and Science University (MSU) Malaysia
Email: dhoooma62@gmail.com, asad_ur@msu.edu.my

DOI Link: http://dx.doi.org/10.6007/IJARPED/v15-i4/29296

Published Online: 13 November 2026

Abstract

Concern has grown that students' reliance on generative artificial intelligence (AI), and its use for academic dishonesty, may weaken self-efficacy, motivation, and academic achievement, yet evidence on these relationships is scattered across inconsistent labels. Following the Arksey and O'Malley framework and the PRISMA extension for scoping reviews (PRISMA-ScR), this review mapped studies published between 2015 and 2025 that examined AI dependence or AI-related academic integrity alongside motivational constructs or academic achievement among higher education students. A Scopus search returned 172 records, of which 21 met the eligibility criteria and were charted. No eligible study predated 2024, eight were set in East Asia and none in Africa, and sixteen were cross-sectional surveys. The apparent inconsistency in the direction of the self-efficacy effect largely disappears once academic self-efficacy, which tended to protect against dependence, is distinguished from AI-specific self-efficacy, which tended to promote it. Dependence was associated with inflated self-efficacy, procrastination, and weaker learning, while associations with achievement were mixed. Amotivation and degree apathy predicted AI-assisted misconduct more strongly than AI use itself. The evidence base remains recent, geographically narrow, reliant on self-report, and almost without longitudinal or intervention research.

Keywords: Generative Artificial Intelligence, Ai Dependence, Self-Efficacy, Academic Integrity, Academic Achievement

Introduction

Generative artificial intelligence (AI) has entered higher education faster than any earlier educational technology. Within months of the public release of ChatGPT in November 2022, students were using large language models to draft essays, explain concepts, debug code, and summarise readings, and universities were revising their assessment and integrity policies in response (Kasneci et al., 2023). Unlike earlier applications such as intelligent tutoring and learning analytics, which supported learners while leaving the cognitive work with them

(Zawacki-Richter et al., 2019), generative tools produce complete artefacts on request and can therefore take over the very work that assessment is designed to capture.

Two risks recur in institutional and scholarly debate. The first is AI dependence: a habitual reliance in which students delegate the drafting, reasoning, problem-solving, or self-monitoring that a task is meant to develop, so that the tool substitutes for, rather than supports, their own effort (Risko & Gilbert, 2016; Zhang et al., 2024). It differs from ordinary use, in which a student consults the tool while remaining responsible for the thinking, and its most visible behavioural form is what Fan et al. (2025) call metacognitive laziness. The second is AI-assisted academic dishonesty, in which generated work is submitted as the student's own (Cotton et al., 2024; Playfoot et al., 2024). The two share a common core, the replacement of the student's effort by the tool, and both have been offered as reasons why the efficiency gains of AI may not translate into learning (Gerlich, 2025).

The phenomenon is wide in scope, because generative tools are freely available, require no technical training, and can be applied to almost any academic task; more than a third of students in one Australian sample had already used chatbots in their assessments, often without regarding this as a breach of integrity (Gruenhagen et al., 2024). Its relevance lies in what a degree certifies. If grades increasingly reflect what a tool produced rather than what a student learned, they lose their value as evidence of achievement, yet universities are writing AI policies with little consolidated evidence on which students become dependent, why, and with what consequences.

These risks are also motivational at root. Self-efficacy, a student's belief in their capability to succeed, is among the strongest psychological correlates of university performance (Bandura, 1997; Richardson et al., 2012). Self-determination theory predicts that amotivated students are more likely to disengage and to outsource effort (Deci & Ryan, 2000), and the Interaction of Person–Affect–Cognition–Execution (I-PACE) model treats problematic technology use as a compensatory response to stress and low perceived competence (Brand et al., 2016). The question, then, is not simply whether students use AI, but how their motivational states shape their reliance on it, and what that reliance in turn does to their motivation and achievement. The central problem addressed by this review is that, although concern about these behaviours is widespread and empirical studies are accumulating rapidly, the evidence on how they relate to students' motivation and achievement is fragmented and, on key points, apparently contradictory. The phenomenon appears under many labels, including "AI dependence", "GenAI reliance", "ChatGPT addiction", "over-reliance", and "metacognitive laziness", each with its own instrument; self-efficacy is measured as academic, task-specific, or AI-specific confidence, often under the same label; and achievement ranges from students' perceptions of their own performance to institutional grades. It therefore remains unclear whether reliance on AI is driven by motivational deficits or by confidence in using the tool, whether dependence weakens self-efficacy and achievement or merely accompanies them, and whether AI-assisted misconduct reflects the availability of the technology or students' disengagement from their studies. Existing reviews of AI in higher education have mapped applications, institutional responses, and the opportunities and challenges of large language models (Kasneci et al., 2023; Zawacki-Richter et al., 2019), but none has examined AI dependence and AI-related integrity together as motivational phenomena or linked them to measured achievement. Without such a synthesis, researchers cannot build cumulatively on

a fast-growing literature, and institutions risk basing policy on assumptions that the evidence has not yet tested.

In response to this problem, this scoping review maps the evidence published between 2015 and 2025 on AI dependence and AI-related academic integrity in relation to motivational constructs and academic achievement among higher education students. Specifically, we address these research questions (RQs):
1. How have AI dependence, AI-related academic integrity, and the associated motivational constructs been conceptualised and measured?
2. What associations have been reported between AI dependence or AI-assisted misconduct, motivational constructs (self-efficacy, motivation, amotivation, and procrastination), and academic achievement?
3. What moderators, protective factors, and interventions have been identified or proposed?
4. What research trends, evidence gaps, and priorities for future research does the evidence reveal?

Methods

This scoping review followed the methodological framework of Arksey and O'Malley (2005), as refined by Levac et al. (2010) and the guidance of the Joanna Briggs Institute (Peters et al., 2020), and is reported according to the PRISMA extension for scoping reviews (PRISMA-ScR; Tricco et al., 2018). A scoping design was chosen because the aim was to map the extent, range, and nature of the evidence in a new and conceptually fragmented field, rather than to estimate a pooled effect. No protocol was registered or published in advance. The objectives, the Population–Concept–Context (PCC) frame, and the eligibility criteria were nevertheless fixed before screening began. Screening and data charting were carried out by the first author and checked by the second author, with disagreements resolved through discussion.

Literature search

Search terms were derived from the PCC frame and organised into three blocks: (a) higher education population terms; (b) AI reliance or AI-related integrity terms; and (c) motivational and achievement terms. Synonyms within each block were combined with OR and the blocks with AND. Within the second block, AI terms were linked to reliance and integrity terms with a proximity operator (W/5), so that a record had to associate AI with dependence or dishonesty rather than mention the two separately. The string was refined once during piloting. An initial version that required motivation and achievement terms jointly, and treated integrity only as an alternative to reliance, returned 11 records and was judged too restrictive; integrity was therefore made a parallel arm, and the motivation and achievement terms were merged into a single block.

The search was run in Scopus in October 2026, with publication years limited to 2015–2025 within the string. Scopus was selected because it is the largest curated abstract and citation database of peer-reviewed literature, with broader journal coverage than Web of Science and strong indexing of the education, psychology, and information-systems literatures on which this topic draws (Baas et al., 2020; Mongeon & Paul-Hus, 2016). The window deliberately opened well before the release of ChatGPT, so that the onset of the literature could be dated rather than assumed. The full search string and filters are provided in Appendix A (Table A1).

Inclusion and Exclusion Criteria

Eligibility was defined through the PCC framework. The population comprised higher education students at undergraduate, postgraduate, or doctoral level; studies with mixed samples were eligible when higher education students formed the main sample. The concept had two arms. The first, AI reliance, was treated as an umbrella construct covering AI dependence or dependency, over-reliance, addiction or problematic use, and substitutive or metacognitively lazy use of AI. The second, AI-related academic integrity, covered AI-assisted plagiarism, cheating, misconduct, and willingness to submit AI-generated work. Either arm had to be examined alongside at least one motivational construct (self-efficacy, motivation, amotivation or apathy, engagement, or procrastination) or academic achievement. The context was open to any country, discipline, or type of institution. Only empirical, peer-reviewed journal articles and conference papers published in English were eligible. Studies of AI adoption or acceptance were excluded unless they also examined reliance or integrity, as were school, teacher, and general-public samples, integrity studies unrelated to AI, and reviews, editorials, and conceptual papers (Table 1).

Table 1
Inclusion and exclusion criteria

Phase/dimension	Criteria
Title/abstract screening (inclusion)	Peer-reviewed journal or conference paper; higher education student population; AI reliance or AI-related academic integrity term; motivational or achievement term; published 2015–2025 in English.
Eligibility assessment (inclusion)	AI reliance or AI-related integrity examined as a studied variable (determinant, outcome, mediator, or moderator), or as a substantive qualitative theme; at least one motivational construct or academic achievement examined; empirical primary study of any design.
Exclusion	AI use, adoption, or acceptance measured without any reliance or integrity construct; school, teacher, employee, or general-public samples; integrity not related to AI; reviews, editorials, and conceptual papers; non-peer-reviewed sources.

Note. Studies of AI adoption were excluded unless they also examined reliance or integrity, because adoption alone does not address the review's concept.

Study selection

Records were screened in two stages. In the first, a keyword-assisted pass over titles, abstracts, and keywords advanced a record only if it contained terms for all four concept groups (students, AI reliance or integrity, motivation, and achievement) and was not a review. This rule was stricter than the eligibility criteria, which require motivation or achievement rather than both. In the second stage, each remaining abstract was assessed against the PCC criteria, and eligible records were sought in full text through publisher, open-access, repository, and author-manuscript routes.

Figure 1 summarises the selection process. The search returned 172 records; because the year limit was applied within the string, none fell outside the window, and no duplicates were found. In the first stage, 123 records were excluded: 114 lacked a motivational or achievement term, 8 were reviews, and 1 lacked an AI reliance or integrity term. In the second stage, a further 28 were excluded because they measured AI use or adoption without a reliance or integrity construct (n = 15), examined no motivational or achievement variable or reported too little to judge (n = 7), sampled an ineligible population (n = 4; one school and

three teacher or academic samples), or addressed integrity unrelated to AI (n = 2). Full text was obtained for 16 of the remaining 21 reports, all of which met the criteria. The other five were not available in full text at the time of charting; they were retained, charted from their abstracts and record information, and are flagged as such in Appendix B.

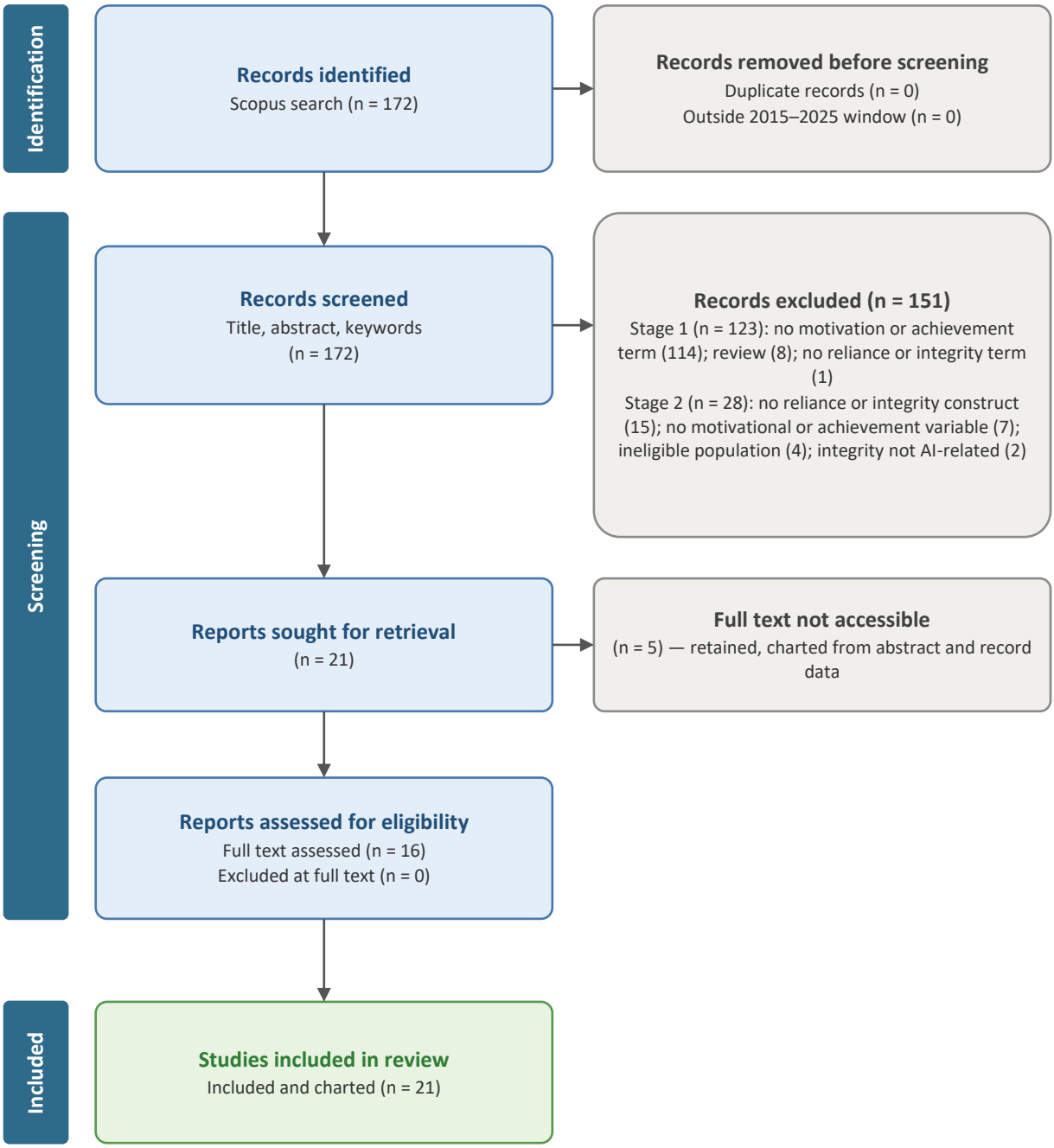


Fig. 1. PRISMA-ScR flow diagram of the study selection process

Coding and data analysis

Data were charted on a template capturing bibliographic details, country, design and analytic method, sample, theoretical framework, the reliance or integrity construct and its measurement, the motivational and achievement constructs, the role and direction of each reported association, moderators and proposed interventions, and author-stated gaps. Constructs were coded in two steps. Construct labels and, where reported, item wording were

first coded inductively into discrete construct types. These types were then grouped into higher-order dimensions according to what each construct describes: the student's reliance on AI, the integrity of the work produced, the student's motivational state, or the academic outcome. The analysis combined descriptive frequency counts with narrative and thematic synthesis.

Four simplifications should be kept in mind when reading the results. Each study was assigned to a single primary evidence strand (dependence, integrity, or both). Publication years follow the Scopus record. Heterogeneously labelled reliance constructs were mapped onto one umbrella construct, which aids comparison but merges sub-constructs that the primary studies treated as distinct. Finally, achievement measured in different ways (self-reported performance, institutional grades, class rank, and task scores) was charted in a common field, so a reported association with achievement does not denote the same outcome across studies.

Quality appraisal

Although critical appraisal is optional in scoping reviews, the included studies were appraised for interpretive purposes with the Mixed Methods Appraisal Tool (MMAT) version 2018 (Hong et al., 2018). Cross-sectional and correlational surveys were appraised under the quantitative non-randomised category, the experiment under the randomised controlled category, and the mixed-methods and qualitative studies under their respective categories. Criterion 3.5 (exposure as intended) was rated "Yes" only when the sample was confirmed to consist of AI users. In line with the developers' advice against overall scores, ratings were recorded for each criterion (available from the authors on request), and the number of criteria met is used only as a descriptor (4–5 higher, 2–3 moderate, 0–1 lower). The appraisal was used to characterise the evidence, not to weight the synthesis or to exclude studies.

Results

Characteristics of included studies

We first summarise the profile of the 21 included studies (Appendix B, Table B1). Although the search window opened in 2015, no eligible study appeared before 2024: three were published in 2024 and eighteen in 2025 (Fig. 2). The field is therefore entirely a product of the generative AI era. Mainland China contributed six studies and Taiwan and South Korea one each, so that eight of the twenty-one were set in East Asia. Romania contributed two, and Peru, Pakistan, Spain, Kuwait, Indonesia, the Philippines, Australia, Canada, and the United Kingdom one each; two studies did not report or confirm a country (Fig. 3). No study was conducted in Africa.

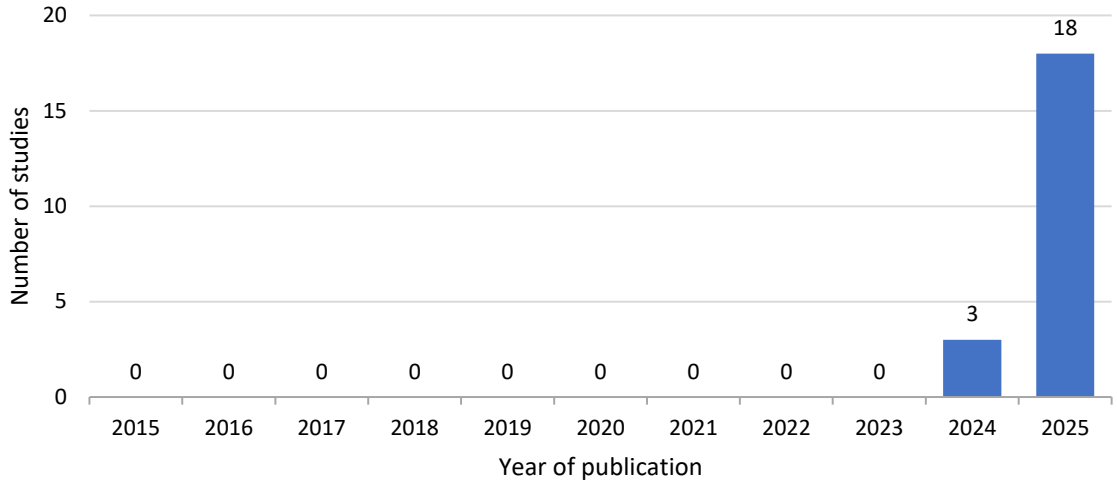


Fig. 2. Included studies by publication year, 2015–2025 (n = 21)

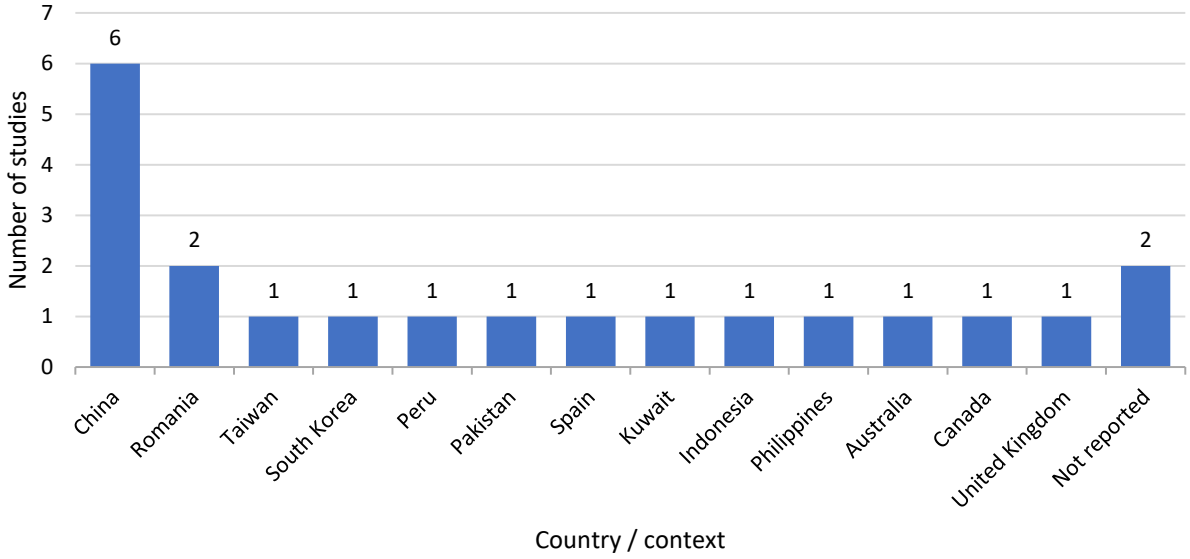


Fig. 3. Included studies by country or context (n = 21)

Methodological quality was modest overall. Of the sixteen studies appraised from full text, two met at least four of the five MMAT criteria (Jia et al., 2025; Wei et al., 2025), seven met two or three, and seven met one or none; the five studies without full text could not be fully appraised. The criterion most often unmet was sample representativeness (3.1), rated “No” in nine of the twelve fully appraised surveys. This reflects a reliance on convenience, volunteer, and single-site samples that limits the generalisability of the mapped findings. Inconsistencies between abstracts, text, and tables were also noted in four studies (Adiyono et al., 2025; Morales-García et al., 2025; Subhani et al., 2025; Vieriu & Petrea, 2025).

Sixteen studies were cross-sectional surveys, most analysed with structural equation, partial least squares, or regression-based mediation models; three used mixed methods (Adiyono et al., 2025; Huang & Wu, 2025; Vieriu & Petrea, 2025), one was a randomised laboratory experiment (Fan et al., 2025), and one was qualitative (Abidi et al., 2025; Fig. 4). None was longitudinal. Twelve studies examined AI dependence or reliance, five examined AI-related integrity, and four addressed both (Fig. 5). Theoretical grounding was thin and dispersed. The

I-PACE model (Ye et al., 2025; Zhang et al., 2024), media system dependency theory (Jia et al., 2025), social cognitive theory (Song & Liu, 2025), the technology acceptance model together with self-determination theory (Dogaru et al., 2025), and UTAUT2 (Subhani et al., 2025) each appeared, while many studies were exploratory or atheoretical.

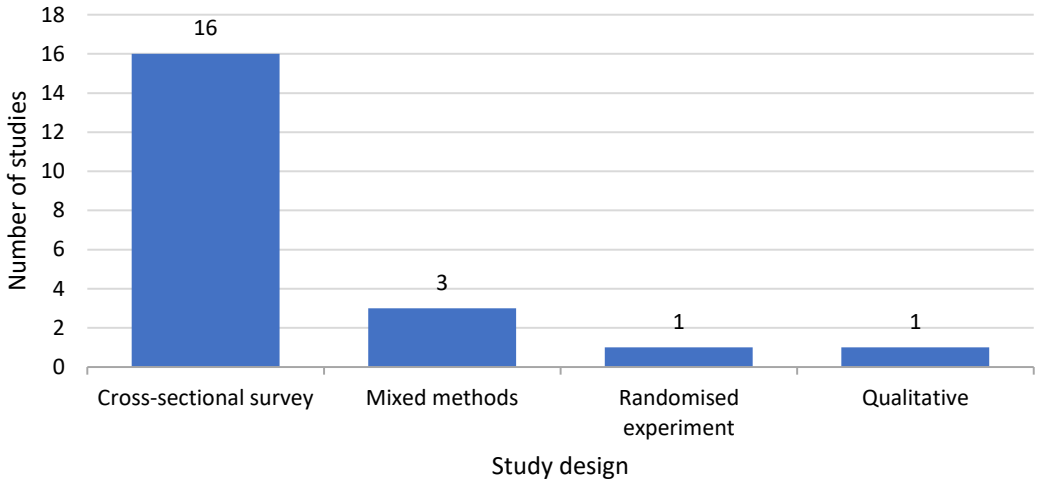


Fig. 4. Included studies by study design (n = 21)

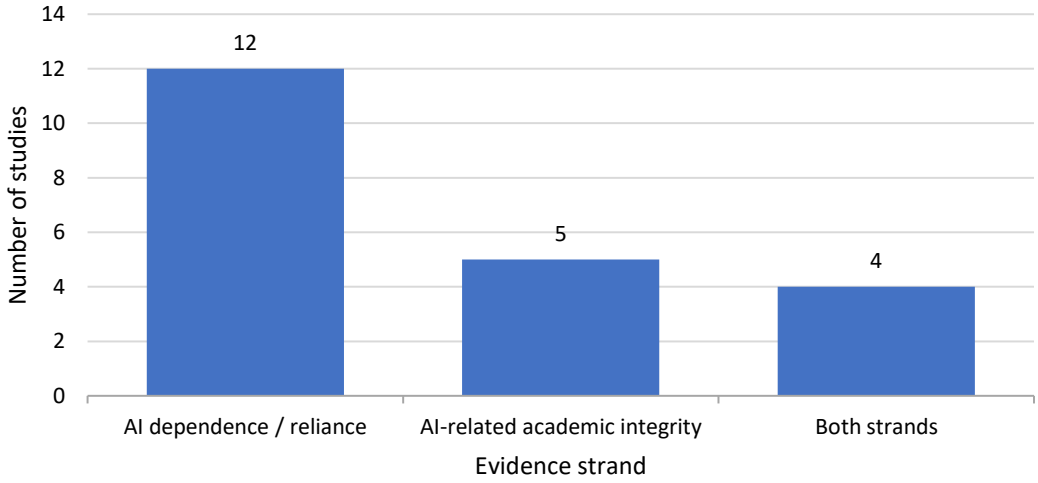


Fig. 5. Included studies by evidence strand (n = 21)

Only 21 of the 172 screened records combined a higher education sample, a reliance or integrity construct, and a motivational or achievement variable. This is a modest evidence base for a phenomenon that now touches most students' academic work. Table 2 cross-tabulates region against evidence strand; blank cells mark combinations for which no evidence was located and therefore indicate where primary research is absent.

Table 2
Coverage matrix of included studies by region × evidence strand (count of studies; blank = no evidence located)

Region	AI dependence / reliance	AI-related integrity	Both strands	Total
East Asia (China, Taiwan, South Korea)	7	1		8
Europe (Romania, Spain, United Kingdom)		2	2	4
South and Southeast Asia (Pakistan, Indonesia, Philippines)	1	1	1	3
Latin America (Peru)	1			1
Middle East (Kuwait)			1	1
Oceania (Australia)		1		1
North America (Canada)	1			1
Not reported	2			2
Total	12	5	4	21

Note. Each study is allocated to one primary evidence strand. Blank cells denote combinations for which no evidence was located.

RQ1: Conceptualisation and measurement of the constructs

AI reliance was conceptualised in at least five ways. The most common was general AI dependence or dependency, measured with short self-report scales adapted from technology-dependence research (Morales-García et al., 2025; Wei et al., 2025; Zhang et al., 2024). A second stream measured dependence on generative AI specifically (Huang & Wu, 2025; Jia et al., 2025), and a third framed it as ChatGPT dependence or addiction, drawing on the problematic-use literature (Lin et al., 2025; Ye et al., 2025). A fourth treated over-reliance as a perceived risk, described or explored qualitatively rather than measured (Abidi et al., 2025; Vieriu & Petrea, 2025). The fifth, and the most distinct, defined reliance behaviourally: "metacognitive laziness" refers to the observed tendency to let the tool carry out the monitoring and revising that learners would otherwise do themselves (Fan et al., 2025). Integrity was operationalised as self-reported plagiarism (Galindo-Domínguez et al., 2025), willingness to submit AI-written work (Playfoot et al., 2024), chatbot use in assessments (Gruenhagen et al., 2024), academic misconduct (Song & Liu, 2025), or integrity as a personal disposition (Subhani et al., 2025). Table 3 consolidates these construct types and the studies that examine each.

Table 3

Constructs identified across the included studies, grouped into four dimensions

Dimension and construct	Description	Evidence (studies)	n
Reliance: general AI dependence	Self-reported reliance on AI tools for academic tasks, measured with short dependence scales.	Zhang et al. (2024); Morales-García et al. (2025); Zheng et al. (2025); Zhang and Xu (2025); Wei et al. (2025); Hernandez et al. (2025)	6
Reliance: GenAI-specific dependency	Dependence on generative AI specifically, often with piloted or adapted scales.	Jia et al. (2025); Huang and Wu (2025)	2
Reliance: ChatGPT dependence or addiction	Problematic-use framing of a single tool.	Ye et al. (2025); Lin et al. (2025); Uppal and Hajian (2025)	3
Reliance: perceived over-reliance	Over-reliance reported as a risk or theme rather than measured.	Vieriu and Petrea (2025); Dogaru et al. (2025); Abidi et al. (2025); Adiyono et al. (2025)	4
Reliance: metacognitive laziness	Observed delegation of self-regulation to the tool during a task.	Fan et al. (2025)	1
Integrity: AI-assisted plagiarism or misconduct	Self-reported plagiarism, misconduct, willingness to submit AI work, or chatbot use in assessments.	Galindo-Domínguez et al. (2025); Playfoot et al. (2024); Gruenhagen et al. (2024); Song and Liu (2025); Subhani et al. (2025)	5
Motivation: academic or learning self-efficacy	Confidence in one's own capability to learn and perform without reference to the tool.	Zhang et al. (2024); Huang and Wu (2025); Jia et al. (2025); Wei et al. (2025); Playfoot et al. (2024); Abidi et al. (2025); Zhang and Xu (2025); Hernandez et al. (2025)	8
Motivation: AI-specific self-efficacy	Confidence in one's capability to operate AI tools.	Morales-García et al. (2025); Lin et al. (2025); Subhani et al. (2025); Zheng et al. (2025)	4
Motivation: intrinsic or learning motivation	Interest, enjoyment, or motivation to learn.	Fan et al. (2025); Adiyono et al. (2025); Dogaru et al. (2025); Gruenhagen et al. (2024); Subhani et al. (2025)	5
Motivation: amotivation, apathy, or avoidance	Lack of perceived value in study, or motivation to avoid effort.	Galindo-Domínguez et al. (2025); Playfoot et al. (2024); Ye et al. (2025)	3
Motivation: procrastination	Delay of academic tasks as a self-regulatory failure.	Uppal and Hajian (2025); Song and Liu (2025)	2
Outcome: academic achievement	Self-reported performance, institutional grades, class rank, or task scores.	Jia et al. (2025); Wei et al. (2025); Fan et al. (2025); Playfoot et al. (2024); Adiyono et al. (2025); Vieriu and Petrea (2025); Uppal and Hajian (2025); Galindo-Domínguez et al. (2025); Abidi et al. (2025)	9

Note. The n column reports the number of included studies that examine each construct; a study may appear under more than one construct, and all twenty-one included studies appear at least once. Per-study construct details are given in Appendix B (Table B1).

The construct types resolve into four higher-order dimensions: reliance on AI, the integrity of the work produced, the student's motivational state, and the academic outcome. Within the motivational dimension, the mapping exposes a distinction that the primary studies rarely make explicit. Academic or learning self-efficacy, the belief that one can succeed through one's own capability, and AI-specific self-efficacy, the belief that one can operate the tool

well, were examined in eight and four studies respectively. Yet the two were seldom measured together and were often reported under the same label. A parallel distinction runs through the reliance dimension, between reliance as a self-reported disposition and reliance as observed behaviour, which only one study measured directly (Fan et al., 2025).

RQ2: Associations with motivation and achievement

The reported associations fall into four groups. The first concerns self-efficacy as an antecedent of dependence. Academic self-efficacy was related to AI dependence mostly indirectly: through academic stress and performance expectations (Zhang et al., 2024), through anxiety and expectations with no direct path (Huang & Wu, 2025), or through emotional pathways (Zheng et al., 2025). AI-specific self-efficacy tended to pull in the opposite direction. It was strongly associated with dependence among Peruvian students ($r = .61$; Morales-García et al., 2025) and rose alongside dependence as use increased (Zhang & Xu, 2025), although one study found ChatGPT self-efficacy associated with lower addiction (Lin et al., 2025). Avoidance learning motivation, fed by inert thinking, was a further antecedent (Ye et al., 2025).

The second group treats dependence as an antecedent. GenAI dependency was associated with inflated, or "false", self-efficacy alongside lower achievement (Jia et al., 2025). In the only experiment, students supported by ChatGPT showed different self-regulatory processes and higher essay scores than comparison groups, but no greater knowledge gain or transfer and no difference in intrinsic motivation (Fan et al., 2025). Dependence also travelled with procrastination, correlating with it at $r = .55$ in one study (Uppal & Hajian, 2025), while procrastination was among the pathways linking dark personality traits to GenAI use habits in another (Song & Liu, 2025).

The third group concerns achievement itself, where the picture was mixed. Nine studies included achievement in some form, mostly self-reported. The negative association came from a study of GenAI dependency and self-reported achievement (Jia et al., 2025). Positive or neutral associations came from a study in which dependency mediated part of a positive effect of AI adoption on class rank (Wei et al., 2025), and from descriptive and mixed-methods studies of perceived gains (Adiyono et al., 2025; Vieriu & Petrea, 2025). Only three studies used objective or institutional outcomes: class rank (Wei et al., 2025), institutional grades (Playfoot et al., 2024), and task and knowledge scores (Fan et al., 2025).

The fourth group concerns integrity, where motivational deficits predicted AI-assisted misconduct more strongly than the frequency of AI use. Amotivation and a cheating culture predicted plagiarism, whereas ChatGPT use alone explained 3.9% of its variance (Galindo-Domínguez et al., 2025). Degree apathy was the only significant predictor of willingness to use AI to write assignments; grades and study-skills self-efficacy were not (Playfoot et al., 2024). More than a third of Australian students had used chatbots in assessments, often without regarding this as a breach (Gruenhagen et al., 2024). Dark personality traits fed academic dishonesty and procrastination (Song & Liu, 2025), and extensive reliance accompanied weaker citation practice (Dogaru et al., 2025). Across these studies, the motivational state that students bring to AI appears to matter at least as much as the tool itself.

RQ3: Moderators, protective factors, and interventions

The studies converge on the view that access to AI matters less than the conditions under which it is used. At the teaching level, perceived teacher caring moderated the link between GenAI dependency and self-efficacy (Jia et al., 2025), and lecturer guidance and learning motivation outweighed AI use in explaining academic success (Adiyono et al., 2025). At the institutional level, organisational AI readiness moderated the effect of AI adoption on performance (Wei et al., 2025), and in a qualitative study a structured policy requiring students to explain and defend their AI-assisted work was associated with confidence, efficiency, and performance (Abidi et al., 2025). At the student level, personal academic integrity moderated whether the intention to use ChatGPT became actual use (Subhani et al., 2025), while academic anxiety and performance expectations emerged as mediators that could be targeted (Huang & Wu, 2025; Zhang et al., 2024). Authors also proposed AI literacy instruction, assessment redesign, and clearer institutional guidance. No study, however, evaluated an intervention designed to reduce dependence or misconduct; the one experiment compared types of learning support (Fan et al., 2025) rather than testing such a measure.

RQ4: Research trends, evidence gaps, and priorities

Within two years the field has moved through compressed stages: from identifying dependence and misconduct as concerns (Gruenhagen et al., 2024; Vieriu & Petrea, 2025), through modelling their motivational antecedents (Morales-García et al., 2025; Ye et al., 2025; Zhang et al., 2024), to the first attempts to link dependence with measured learning (Fan et al., 2025; Jia et al., 2025; Wei et al., 2025). The evidence is stronger in describing the correlates of dependence than in establishing its consequences or evaluating responses to it. Table 4 charts the main gaps, together with a recommended design and outcome measure for each: the recency of the evidence, its concentration in East Asia and absence from Africa, the dominance of cross-sectional designs, reliance on self-reported achievement, inconsistent measurement of dependence and self-efficacy, convenience sampling, and the lack of evaluated interventions.

Table 4

Evidence gaps charted from the included studies, with a recommended study design and outcome measure for each

Gap	Evidence from mapping	Recommended design	Outcome measure
Temporal recency	21 of 21 studies from 2024–2025	Updating reviews and replication of early findings	Stability of associations over time
Geographic concentration	8 of 21 from East Asia; none from Africa	Primary and multi-country studies in Africa, the Gulf, and Latin America	Cross-cultural comparability of constructs
Design limitations	16 of 21 cross-sectional; none longitudinal	Panel designs over at least one semester; cross-lagged models	Direction of self-efficacy–dependence effects
Outcome limitations	9 measured achievement; 3 used objective measures	Linkage to institutional records and transfer tasks	GPA, examination scores, unassisted transfer performance
Construct fragmentation	Five reliance framings; academic and AI-specific self-efficacy conflated	Instrument development and validation	Validated scales measuring both forms of self-efficacy

Gap	Evidence from mapping	Recommended design	Outcome measure
Measurement quality	Single-item and ad hoc scales in several studies	Psychometric validation with confirmatory factor analysis	Reliability and measurement invariance
Intervention evidence	No evaluated intervention	Randomised and quasi-experimental trials	Change in dependence, misconduct, and learning

Note. Gaps are charted from the author-stated limitations of the included studies together with the authors' cross-study mapping.

Discussion and Conclusion

This scoping review mapped what is known about AI dependence and AI-related academic integrity in relation to motivation and academic achievement among higher education students. Across 21 studies, all published in 2024 or 2025, the literature treats AI dependence less as a technological issue than as a motivational and self-regulatory one. These findings should be read as patterns in a young and largely cross-sectional literature, not as causal claims. Because a scoping review neither pools effects nor weights studies by quality, what follows concerns where the evidence has accumulated and how it may be interpreted, rather than how large the effects are (Arksey & O'Malley, 2005; Levac et al., 2010; Tricco et al., 2018).

Motivation and reliance as a two-way relationship

Research on educational technology has generally treated motivation as a predictor of use. The mapped studies suggest that, for generative AI, the relationship runs both ways. Low academic confidence, stress, and avoidance motivation predicted dependence (Huang & Wu, 2025; Ye et al., 2025; Zhang et al., 2024), and dependence in turn was associated with inflated confidence, procrastination, and weaker learning (Fan et al., 2025; Jia et al., 2025; Uppal & Hajian, 2025). If confirmed longitudinally, such a cycle would explain why dependence may become self-reinforcing.

These patterns sit comfortably within literatures well beyond AI in education. The compensatory route from stress and low perceived competence to reliance is what the I-PACE model predicts for problematic technology use (Brand et al., 2016). The delegation of monitoring and revision to the tool fits research on cognitive offloading, in which external aids reduce effort at the expense of internal processing (Gerlich, 2025; Risko & Gilbert, 2016), and the link with procrastination echoes the meta-analytic account of procrastination as a failure of self-regulation (Steel, 2007). The role of AI-specific self-efficacy in driving use is what technology acceptance models would predict (Davis, 1989; Venkatesh et al., 2003). In the integrity strand, the predictive weight of amotivation and apathy is consistent with self-determination theory (Deci & Ryan, 2000), and moral disengagement theory offers a well-established account of how students may excuse such rule-breaking to themselves (Bandura, 1999).

Two kinds of self-efficacy

The clearest disagreement in the corpus concerns the direction of the self-efficacy effect. Some studies found higher self-efficacy accompanying lower dependence (Huang & Wu, 2025; Lin et al., 2025; Zhang et al., 2024), others found it accompanying higher dependence (Morales-García et al., 2025; Zhang & Xu, 2025), and one found that dependence raised self-

efficacy while lowering achievement (Jia et al., 2025). We suggest that much of this disagreement is a matter of measurement rather than substance. Academic self-efficacy, confidence in one's own capability to learn and perform, tends to protect against dependence, because confident students have less need to compensate. AI-specific self-efficacy, confidence in operating the tool, tends to encourage dependence, because skill and ease of use invite further use. A third form, which we term borrowed self-efficacy, arises when success achieved with AI is credited to oneself. This is the inflated self-efficacy reported by Jia et al. (2025), and it resembles the miscalibration between confidence and competence documented in the wider literature (Kruger & Dunning, 1999).

Once these forms are held apart, the protective effect of the first and the promoting effect of the second become complementary rather than contradictory, and the third explains how dependence can coexist with high reported confidence. One finding does not fit neatly: ChatGPT self-efficacy was associated with lower addiction in one study (Lin et al., 2025). This may reflect the difference between dependence and addiction framings, since skilled and confident users may use a tool heavily without the loss of control that addiction scales capture, but the point warrants direct testing. The practical implication concerns measurement. Future instruments should measure academic and AI-specific self-efficacy separately, and should assess academic self-efficacy for tasks completed without AI, so that borrowed confidence can be distinguished from genuine capability.

Performance is not the same as learning

A related tension concerns achievement. Studies of AI adoption reported positive associations with performance (Vieriu & Petrea, 2025; Wei et al., 2025), whereas studies of dependence reported negative ones (Jia et al., 2025). The two are reconcilable once adoption and dependence are understood as different patterns of use. Use that supports a student's own reasoning may raise performance, whereas use that substitutes for it may raise task scores while leaving learning unchanged. The only study that measured outcomes experimentally found precisely this dissociation, with better essays but no greater knowledge gain or transfer (Fan et al., 2025). Studies that measure performance on AI-assisted tasks, or students' perceptions of their own performance, may therefore record gains that would not survive the removal of the tool. The relationship between AI use and achievement is better modelled as conditional on the pattern of use than as a single linear effect.

Integrity as a motivational question

Two developments distinguish this literature from earlier research on technology in learning. The first is the emergence of a behavioural, rather than attitudinal, measure of reliance. The observation of metacognitive laziness in a controlled setting (Fan et al., 2025) shows that dependence can be studied in what students do, not only in what they report. The second is the finding that AI-assisted misconduct is driven more by motivation than by access. Two studies independently found that amotivation and degree apathy, rather than the frequency or availability of AI, predicted plagiarism and willingness to submit AI-written work (Galindo-Domínguez et al., 2025; Playfoot et al., 2024). This moves the analytical focus from detecting AI use to understanding why some students choose to misuse it. Both developments point to AI dependence being reconceived as a motivational and self-regulatory phenomenon, and the four-dimension framework proposed here is intended to support that reconceptualisation.

Implications for Research

Several implications follow for future research. First, the direction of the self-efficacy–dependence relationship can only be established through longitudinal and cross-lagged designs that follow students across at least one semester. Second, reliance measures should be linked to objective achievement, such as institutional grades and examination scores, and to performance on unassisted transfer tasks, so that the gap between performance and learning can be tested directly. Third, the field needs validated instruments that measure academic and AI-specific self-efficacy separately and distinguish appropriate use from over-reliance, since several included studies relied on single-item or ad hoc scales. Fourth, research should extend well beyond East Asia, particularly to Africa and the Gulf, where eligible studies were absent or scarce. Finally, interventions, from assessment redesign to programmes that build academic self-efficacy, should be evaluated through randomised and quasi-experimental trials.

Implications for Practice and Policy

For practice, the recurring message is that the motivational conditions of use matter more than the tool. The evidence points toward building academic self-efficacy through mastery experiences completed without AI, so that confidence rests on genuine capability; designing assessments in which students explain, justify, and defend their AI-assisted work (Abidi et al., 2025); and assessing learning in ways the tool cannot complete alone, so that the gap between performance and learning is exposed rather than rewarded. Teacher caring and lecturer guidance appeared protective (Adiyono et al., 2025; Jia et al., 2025) and are low-cost levers available to every instructor.

For policy, the findings support a shift from detection toward motivation and provision. Because AI-assisted misconduct was predicted more by apathy and amotivation than by AI use (Galindo-Domínguez et al., 2025; Playfoot et al., 2024), policies that rely solely on detection and sanction are unlikely to reach its causes. Institutions should pair clear rules on acceptable use with organisational readiness, staff development, and AI literacy provision (Wei et al., 2025), and should treat students' engagement with their studies as part of their integrity strategy. Given the geographic concentration of the evidence, these directions are best grounded for the settings studied and should be adapted and evaluated as they are extended elsewhere.

Limitations and Concluding Remarks

Several limitations should be acknowledged. First, as a scoping review, this study mapped the evidence without pooling results, and the MMAT appraisal was used only for interpretation; the review therefore cannot judge which reported relationships are the most robust. Second, the search was restricted to a single database and to peer-reviewed English-language sources, with no backward or forward citation searching. This may have excluded grey and non-English literature and probably compounded the geographic imbalance documented here. Third, the keyword rule in the first screening stage was stricter than the eligibility criteria and may have excluded eligible studies. Fourth, five included studies were charted from their abstracts; removing them would reduce the evidence on the amotivation–misconduct association to a single study (Playfoot et al., 2024) but would not alter any other theme. Fifth, screening agreement was not quantified, the protocol was not registered, and the grouping of constructs into four dimensions is interpretive. Finally, the 2015–2025 window excludes a

fast-growing 2026 literature; a preliminary search identified more than fifty potentially eligible 2026 records, so an update will soon be warranted. Because the underlying studies rely largely on cross-sectional self-report, the review also inherits that limitation: it maps what students report about their reliance and achievement rather than what they do or learn.

This review shows that research on AI dependence in higher education, although recent, already points in a consistent direction. Weak academic confidence, stress, and avoidance motivation are associated with greater reliance on AI, and reliance is associated with inflated confidence, procrastination, and task gains that do not carry over into learning. AI-assisted misconduct, in turn, appears to be rooted more in amotivation and apathy than in the availability of the tools. The review bounds a fragmented field within four dimensions, shows that the apparent disagreement over self-efficacy largely dissolves once academic, AI-specific, and borrowed self-efficacy are distinguished, and charts where evidence is concentrated and where it is missing.

The question facing the field is therefore no longer whether students rely on AI, but under what motivational and instructional conditions that reliance supports learning rather than replacing it. Answering it will require longitudinal designs, objective outcomes, validated measures, wider geographic coverage, and evaluated interventions. By clarifying current patterns and gaps, this review offers a foundation for subsequent systematic reviews, primary research, and evidence-informed policy in higher education.

Author contributions
Conceptualization, A.O. and A.U.; methodology, A.O. and A.U.; validation, A.U.; formal analysis, A.O.; investigation, A.O.; data curation, A.O.; writing – original draft preparation, A.O.; writing – review and editing, A.U.; visualization, A.O.; supervision, A.U.; project administration, A.U. All authors read and approved the final manuscript.

Funding
This research received no external funding.

Data availability
All data are drawn from the published studies cited and charted in Appendix B; the search strategy is given in Appendix A

Declarations
Competing interests
The authors declare no competing interests.

Use of generative AI
Generative AI (Claude, Anthropic) was used to help retrieve full texts and to edit and condense the authors' writing, and Grammarly was used for language checking. All research decisions were made by the authors, who take full responsibility for the manuscript..

References

Abidi, O., Dzenopoljac, V., & Dzenopoljac, A. (2025). The AI-powered learning loop in higher education. Learning and Teaching in Higher Education: Gulf Perspectives, 19(2), 69–82. https://doi.org/10.1108/LTHE-03-2025-0014

Adiyono, A., Suwartono, T., Nurhayati, S., Dalimarta, F. F., & Wijayanti, O. (2025). Impact of artificial intelligence on student reliance for exam answers: A case study in IRCT Indonesia. International Journal of Learning, Teaching and Educational Research, 24(3), 455–479. https://doi.org/10.26803/ijlter.24.3.22

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. International Journal of Social Research Methodology, 8(1), 19–32. https://doi.org/10.1080/1364557032000119616

Baas, J., Schotten, M., Plume, A., Côté, G., & Karimi, R. (2020). Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. Quantitative Science Studies, 1(1), 377–386. https://doi.org/10.1162/qss_a_00019

Bandura, A. (1997). Self-efficacy: The exercise of control. W. H. Freeman.

Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. Personality and Social Psychology Review, 3(3), 193–209. https://doi.org/10.1207/s15327957pspr0303_3

Brand, M., Young, K. S., Laier, C., Wölfling, K., & Potenza, M. N. (2016). Integrating psychological and neurobiological considerations regarding the development and maintenance of specific Internet-use disorders: An Interaction of Person-Affect-Cognition-Execution (I-PACE) model. Neuroscience & Biobehavioral Reviews, 71, 252–266. https://doi.org/10.1016/j.neubiorev.2016.08.033

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. Innovations in Education and Teaching International, 61(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340. https://doi.org/10.2307/249008

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. Psychological Inquiry, 11(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

Dogaru, M., Pisică, O., Popa, C.-Ș., Răgman, A.-A., & Tololoi, I.-R. (2025). The perceived impact of artificial intelligence on academic learning. Frontiers in Artificial Intelligence, 8, Article 1611183. https://doi.org/10.3389/frai.2025.1611183

Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., Shen, Y., Li, X., & Gašević, D. (2025). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. British Journal of Educational Technology, 56(2), 489–530. https://doi.org/10.1111/bjet.13544

Galindo-Domínguez, H., Campo, L., Delgado, N., & Sainz de la Maza, M. (2025). Relationship between the use of ChatGPT for academic purposes and plagiarism: The influence of student-related variables on cheating behavior. Interactive Learning Environments, 33(6), 4047–4061. https://doi.org/10.1080/10494820.2025.2457351

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

Gruenhagen, J. H., Sinclair, P. M., Carroll, J.-A., Baker, P. R. A., Wilson, A., & Demant, D. (2024). The rapid rise of generative AI and its implications for academic integrity: Students'

perceptions and use of chatbots for assistance with assessments. Computers and Education: Artificial Intelligence, 7, Article 100273. https://doi.org/10.1016/j.caeai.2024.100273

Hernandez, H. E., Malimata, C. A., Lucas, J. D., Pampo, F. R., Grume, J. C., Yambao, J. A., Austria, J. A. P., Salenga, J. L., & Miranda, J. P. P. (2025). Computing students' attitudes toward ChatGPT across learning modalities in programming courses. In 2025 13th International Conference on Information and Education Technology (ICIET) (pp. 44–49). IEEE. https://doi.org/10.1109/ICIET66371.2025.11046273

Hong, Q. N., Fàbregues, 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

Huang, T., & Wu, C. (2025). The chain mediating effect of academic anxiety and performance expectations between academic self-efficacy and generative AI reliance. Computers and Education Open, 9, Article 100275. https://doi.org/10.1016/j.caeo.2025.100275

Jia, W., Pan, L., & Neary, S. (2025). Effect of GenAI dependency on university students' academic achievement: The mediating role of self-efficacy and moderating role of perceived teacher caring. Behavioral Sciences, 15(10), Article 1348. https://doi.org/10.3390/bs15101348

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, 103, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. Journal of Personality and Social Psychology, 77(6), 1121–1134. https://doi.org/10.1037/0022-3514.77.6.1121

Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: Advancing the methodology. Implementation Science, 5(1), Article 69. https://doi.org/10.1186/1748-5908-5-69

Lin, M.-Y., Hong, J.-C., Tai, K.-H., & Jiang, P.-H. (2025). ChatGPT addiction predicted by openness to using new technology and self-efficacy, and reflected in continuous intention. Interactive Learning Environments, 33(8), 4805–4819. https://doi.org/10.1080/10494820.2025.2471910

Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. Scientometrics, 106(1), 213–228. https://doi.org/10.1007/s11192-015-1765-5

Morales-García, W. C., Sairitupa-Sanchez, L. Z., Flores-Paredes, A., Pascual-Mariño, J., & Morales-García, M. (2025). Influence of self-efficacy in the use of artificial intelligence (AI) and anxiety toward AI use on AI dependence among Peruvian university students. Data and Metadata, 4, Article 210. https://doi.org/10.56294/dm2025210

Peters, M. D. J., Marnie, C., Tricco, A. C., Pollock, D., Munn, Z., Alexander, L., McInerney, P., Godfrey, C. M., & Khalil, H. (2020). Updated methodological guidance for the conduct of scoping reviews. JBI Evidence Synthesis, 18(10), 2119–2126. https://doi.org/10.11124/JBIES-20-00167

Playfoot, D., Quigley, M., & Thomas, A. G. (2024). Hey ChatGPT, give me a title for a paper about degree apathy and student use of AI for assignment writing. The Internet and Higher Education, 62, Article 100950. https://doi.org/10.1016/j.iheduc.2024.100950

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. Psychological Bulletin, 138(2), 353–387. https://doi.org/10.1037/a0026838

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. Trends in Cognitive Sciences, 20(9), 676–688. https://doi.org/10.1016/j.tics.2016.07.002

Song, J., & Liu, S. (2025). Dark personality traits are associated with academic misconduct, frustration, negative thinking, and generative AI use habits: The case of Sichuan art universities. BMC Psychology, 13(1), Article 633. https://doi.org/10.1186/s40359-025-02949-w

Steel, P. (2007). The nature of procrastination: A meta-analytic and theoretical review of quintessential self-regulatory failure. Psychological Bulletin, 133(1), 65–94. https://doi.org/10.1037/0033-2909.133.1.65

Subhani, F., Khan, S. A., Sandhu, M. A., & Shahzad, M. F. (2025). What factors affect the adoption intention and actual use of ChatGPT in higher education? The moderating role of academic integrity. TechTrends, 69(5), 1056–1071. https://doi.org/10.1007/s11528-025-01096-8

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., … Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. Annals of Internal Medicine, 169(7), 467–473. https://doi.org/10.7326/M18-0850

Uppal, K., & Hajian, S. (2025). Students' perceptions of ChatGPT in higher education: A study of academic enhancement, procrastination, and ethical concerns. European Journal of Educational Research, 14(1), 199–211. https://doi.org/10.12973/eu-jer.14.1.199

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS Quarterly, 27(3), 425–478. https://doi.org/10.2307/30036540

Vieriu, A. M., & Petrea, G. (2025). The impact of artificial intelligence (AI) on students' academic development. Education Sciences, 15(3), Article 343. https://doi.org/10.3390/educsci15030343

Wei, L., Qiu, L., & Man, S. (2025). The dual impact of AI adoption on college students' academic performance in music education: Evidence from China. SAGE Open, 15(4). https://doi.org/10.1177/21582440251407506

Ye, J.-H., Zhang, M., Nong, W., Wang, L., & Yang, X. (2025). The relationship between inert thinking and ChatGPT dependence: An I-PACE model perspective. Education and Information Technologies, 30(3), 3885–3909. https://doi.org/10.1007/s10639-024-12966-8

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? International Journal of Educational Technology in Higher Education, 16(1), Article 39. https://doi.org/10.1186/s41239-019-0171-0

Zhang, L., & Xu, J. (2025). The paradox of self-efficacy and technological dependence: Unraveling generative AI's impact on university students' task completion. The Internet

and Higher Education, 65, Article 100978. https://doi.org/10.1016/j.iheduc.2024.100978

Zhang, S., Zhao, X., Zhou, T., & Kim, J. H. (2024). Do you have AI dependency? The roles of academic self-efficacy, academic stress, and performance expectations on problematic AI usage behavior. International Journal of Educational Technology in Higher Education, 21(1), Article 34. https://doi.org/10.1186/s41239-024-00467-0

Zheng, H., Hu, X., & Wang, Y. (2025). Beyond usefulness: A cognitive-affective-behavioral analysis of student trust and dependence on generative AI in higher education. In Proceedings of the 2025 International Conference on Educational Technology and Artificial Intelligence (pp. 230–235). ACM. https://doi.org/10.1145/3766557.3766596