



Impact of Artificial Intelligence on Student Autonomy and Identity: A Narrative Literature Review

Saeeda Khoso¹, Junaid Ali Sahito², Samia Saif³

¹ MPhil Scholar, Sukkur IBA University.

Email: saeeda.mphils25@iba-suk.edu.pk

² MPhil Scholar, Lahore University of Management Sciences.

Email: 20170044@lums.edu.pk

³ MPhil Scholar, Sukkur IBA University.

Email: samia.mphils25@iba-suk.edu

ARTICLE INFO

Submission:

August 03, 2026

Revised:

August 25, 2026

Accepted:

September 07, 2026

Available Online:

September 19, 2026

Keywords:

Artificial Intelligence, Student Autonomy, Student Identity, Self-Directed Learning, AI in Education, Learner Agency

Corresponding author:

Saeeda Khoso

Email:

saeeda.mphils25@ibasuk.edu.pk

ABSTRACT

The rapid proliferation of artificial intelligence (AI) tools in educational settings has catalyzed significant scholarly interest in how these technologies reshape two fundamental dimensions of learner development: student autonomy and student identity. This narrative literature review synthesizes evidence from 30 peer-reviewed and scholarly works published predominantly between 2024 and 2026, examining the multifaceted ways in which AI-mediated learning environments influence learners' capacity for self-direction, independent decision-making, and the formation of their academic and personal identities. Drawing on studies conducted across diverse educational contexts (including higher education, secondary schooling, language learning, and creative disciplines), the review reveals a dual and often paradoxical relationship between AI integration and student development. On one hand, AI tools such as adaptive learning platforms, AI-powered feedback systems, and generative models like ChatGPT have been found to enhance student autonomy by offering personalized learning pathways, immediate feedback, and opportunities for self-regulated exploration. On the other hand, critical concerns have emerged regarding learner over-reliance, diminished critical thinking, erosion of epistemic agency, and the destabilization of learner identity when students navigate AI-assisted environments. The review further identifies tensions between productivity gains and identity coherence, and highlights the mediating roles of AI literacy, self-efficacy, and institutional context. Implications for educators, curriculum designers, and policymakers are discussed, alongside recommendations for balanced and ethically informed AI integration in education.

Introduction

The emergence of artificial intelligence as a pervasive force in education marks one of the most consequential transformations in the history of formal learning. From adaptive tutoring systems to large language models capable of generating essays, solving problems, and simulating dialogue, AI technologies have fundamentally altered the environment in which students learn, think, and develop their sense of self. As Wang et al. (2024) observe in their systematic literature review, AI in education encompasses a broad spectrum of applications, ranging from intelligent tutoring systems and automated assessment tools to personalized recommendation engines, each carrying distinct implications for how students engage with knowledge and with their own learning processes.

At the centre of contemporary educational philosophy lies the concept of student autonomy: the capacity of learners to exercise informed, self-directed control over their educational experiences. Rooted in philosophical traditions traced back to Dearden (2010) and operationalized extensively in the work of Boud (2012), learner autonomy encompasses not only the practical ability

to manage one's own learning but also the deeper disposition toward intellectual independence, self-reflection, and the capacity to set, pursue, and evaluate personal learning goals. These capacities, long considered foundational to high-quality education, are now being tested and renegotiated in the presence of AI systems that can automate significant portions of the learning process.

Equally significant is the question of student identity, which refers to the complex and evolving sense of who one is as a learner, thinker, and social being. As Cinoğlu and Arıkan (2012) articulate in their theoretical account, identity formation is a multidimensional process shaped by social interaction, cognitive development, and the environments within which individuals find themselves embedded. When those environments are increasingly mediated by AI systems, the question of how students come to understand themselves, including their strengths, their creative capacities, and their intellectual ownership, becomes urgent and non-trivial.

The purpose of this review is to synthesize the growing body of scholarship examining the impact of AI on student autonomy and student identity. By engaging with 30 carefully selected studies spanning educational levels, disciplinary contexts, and methodological approaches, the review seeks to construct a coherent and critical account of what the current literature reveals about these two interrelated dimensions of student development. The review is guided by three overarching questions: (1) In what ways do AI tools enhance or constrain student autonomy? (2) How do AI-mediated environments influence the formation and expression of student identity? (3) What factors mediate or moderate these effects, and what are the implications for educational practice?

The significance of this inquiry extends beyond academic interest. Desjardins (2015) has argued that education plays a precarious but vital role in identity and value formation, particularly as market forces and technological change reshape educational priorities. In this context, understanding the human developmental consequences of AI integration is not merely a theoretical concern but a pressing practical imperative for educators, institutions, and policymakers navigating an increasingly AI-saturated educational landscape.

Methodology

Review Design

This study employs a narrative literature review methodology, chosen for its capacity to synthesize a heterogeneous body of literature across multiple disciplines, contexts, and methodological traditions. Unlike systematic reviews that prioritize exhaustive search protocols and quantitative synthesis, narrative reviews are particularly well-suited to exploratory and conceptually complex topics where the goal is interpretive integration rather than statistical aggregation (Baumeister & Leary, 1997). Given that the intersection of AI, student autonomy, and student identity spans educational psychology, technology studies, language education, and philosophical inquiry, the narrative approach offers the analytical flexibility required to do justice to the breadth and depth of existing scholarship.

PRISMA-Informed Search and Selection Process

Although this is a narrative review, the selection of sources was guided by a PRISMA-informed (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) approach to ensure transparency and rigor. As Torraco (2005) notes, even integrative and narrative reviews benefit from structured and transparent search procedures that document inclusion and exclusion decisions systematically. Literature searches were conducted across multiple academic databases including Google Scholar, Scopus, Web of Science, and ResearchGate, using search terms including: 'artificial intelligence and student autonomy', 'generative AI and learner identity', 'ChatGPT and self-directed learning', 'AI and learner agency', 'AI and student identity formation', and combinations thereof. The temporal scope was deliberately extended to include foundational theoretical texts alongside the most recent empirical contributions.

Inclusion criteria required that studies (a) addressed AI or generative AI tools in educational contexts, (b) explicitly examined outcomes related to student autonomy, learner agency, or student identity, (c) were published in peer-reviewed journals, conference proceedings, or as scholarly monographs, and (d) were written in English. Studies were excluded if they addressed AI only in non-educational contexts, focused exclusively on teacher perspectives without student dimensions, or lacked sufficient methodological transparency.

Table 1 below presents a summary of the PRISMA-informed search and selection process.

Table 1: PRISMA-Informed Search and Selection Summary

Database/Source	Initial Records	After Screening	Included
Google Scholar	65	38	14
Scopus	48	27	9
Web of Science	32	18	5
ResearchGate/Others	21	12	2
Total	166	95	30

Following the selection process, 30 sources were retained for analysis. These include empirical studies employing qualitative, quantitative, and mixed methods designs; theoretical and conceptual papers; meta-analyses; and one book chapter. The studies span contexts including higher education in the Global South and Global North, secondary English education, EFL (English as a Foreign Language) classrooms, creative graduate programmes, language teacher education, and general school settings.

Analytical Approach

The retained sources were read and re-read using a thematic approach to narrative synthesis. Key themes were identified inductively from the data, then organized deductively against the two primary constructs of student autonomy and student identity. Where studies addressed both constructs or their intersection, they were analyzed under multiple thematic categories. The analysis attends particularly to patterns of convergence and divergence across studies, disciplinary and contextual differences in findings, and the mediating or moderating variables that shape AI's educational effects.

Literature Review

Theoretical Foundations: Autonomy, Identity, and Education

Before examining the empirical literature, it is necessary to establish the conceptual foundations upon which this review rests. The notion of learner autonomy has a rich philosophical lineage. Dearden (2010) situates autonomy within liberal educational philosophy, arguing that education's fundamental purpose is to cultivate individuals capable of reasoning independently, resisting external manipulation, and making authentic choices about their own lives and learning. Boud (2012) extends this philosophical perspective into practical pedagogy, conceptualizing student autonomy as a cluster of capacities including self-assessment, goal-setting, reflection, and the ability to learn beyond formal instruction. For Boud, autonomy is not a fixed state but a developmental achievement nurtured through educational environments that grant students meaningful agency.

Student identity, as theorized by Cinoğlu and Arıkan (2012), is equally complex. Drawing on psychodynamic, social cognitive, and sociocultural perspectives, they describe identity as a dynamic construct shaped by the interplay of self-perception, social interaction, and environmental context. Identity formation is neither linear nor complete; it is a continuous process of negotiation between the individual and the social world. In educational contexts, this process is profoundly influenced by how students perceive their competence, their relationships with teachers and peers, and the cultural norms that govern learning environments. Desjardins (2015) adds a structural dimension, noting that shifts from state to market governance in education have increasingly commodified the educational experience in ways that can destabilize identity formation and reduce the space for authentic value development.

These foundational perspectives provide the lens through which the empirical literature on AI's educational impact can be critically assessed. The question is not simply whether AI tools enhance or diminish measurable learning outcomes, but whether they support or undermine the deeper developmental processes of autonomous learning and identity formation.

AI and Student Autonomy: Enhancement and Enablement

A significant strand of the literature reports positive effects of AI on student autonomy, particularly through personalization, adaptive feedback, and the removal of barriers to independent learning. Mulyani and Hermita (2025) investigate AI-based adaptive learning models and their influence on learning personalization and student autonomy, finding that these systems enable learners to progress at their own pace, receive content calibrated to their current level, and make choices about their learning pathways in ways that traditional classrooms rarely permit. Their findings suggest that when AI is thoughtfully designed to scaffold rather than replace student decision-making, it can meaningfully extend learners' autonomous capacities.

Huma et al. (2025) similarly find that AI-driven feedback systems have a constructive impact on student autonomy and self-directed learning. In their study, students who received personalized, immediate feedback from AI systems demonstrated greater capacity for self-assessment and were more likely to initiate revision and reflective activity than those in traditional feedback conditions. The authors argue that the immediacy and specificity of AI feedback creates conditions favorable to the development of metacognitive habits associated with autonomous learning.

Wang et al. (2026) provide robust meta-analytic support for these findings. Their meta-analysis of AI-supported personalized feedback across multiple studies concludes that such interventions positively influence students' learning outcomes and motivation, with effects extending to dimensions of learner self-regulation and intrinsic engagement. The effectiveness, they note, is moderated by the quality of feedback design, learner characteristics, and the extent to which AI tools are integrated into broader pedagogical frameworks rather than deployed in isolation.

Ng et al. (2024) offer an important empirical contribution from the domain of science education, demonstrating that ChatGPT can empower student self-regulated learning when used in structured, pedagogically intentional ways. Their pioneering pilot study found that secondary students who used ChatGPT as a learning partner, asking questions, checking their understanding, and exploring concepts beyond the curriculum, exhibited enhanced self-regulation and greater ownership of their learning processes. These findings underscore the potential of generative AI as a tool for extending the autonomous reach of learners, provided it is deployed in educationally purposeful contexts.

Almohesh (2024) examines ChatGPT's role in the autonomy of Saudi Arabian primary school students in online classes, reporting that both students and teachers perceived AI as expanding learners' capacity to seek information, pursue questions independently, and engage with English language learning beyond class time. These perceptions align with Kundu and Bej's (2025) study of AI innovations in schools of the Global South, which found that AI tools showed particular promise in empowering EFL students in resource-constrained settings by providing access to individualized language support previously unavailable in their educational contexts.

Mammadova (2026) advances the theoretical argument that AI literacy, understood as the ability to critically understand, evaluate, and use AI tools, is itself a form of learner autonomy in the contemporary world. In her analysis, AI literacy redefines what it means for a student to be autonomous in the era of generative AI, extending the concept beyond traditional self-directed learning to encompass the capacity to critically engage with, rather than passively consume, AI-generated content. This conceptual expansion has significant implications for how educational institutions understand and cultivate learner independence.

Ouchene and Belmekki (2025) explore student attitudes toward AI tools in the learning process with specific attention to motivation and autonomy. Their findings reveal that students who hold positive attitudes toward AI are more likely to experience increased intrinsic motivation and autonomy-supportive learning behaviours. The authors suggest that the motivational dimension of AI use is crucial: when students experience AI as a tool that expands their capacities rather than one that performs for them, the autonomy-enhancing potential is maximized.

The work of Mncube et al. (2026) in South African higher education contexts reinforces this optimistic picture in relation to self-directed learning. They find that AI tools, when appropriately scaffolded within institutional frameworks, support self-directed learning and student autonomy by enabling students to identify their own knowledge gaps, set personalized learning goals, and access resources tailored to their needs. Their study suggests that institutional commitment to pedagogical support is a critical condition for AI's autonomy-enhancing effects to be realized.

AI and Student Autonomy: Risks, Constraints, and Over-Reliance

While the literature documents genuine autonomy-enhancing effects, an equally substantial body of scholarship raises serious concerns about the ways in which AI integration may constrain, diminish, or distort student autonomy. These concerns cluster around three principal themes: over-reliance and passivity, diminished critical thinking, and the undermining of epistemic agency.

Alm (2024) provides a nuanced exploration of autonomy in what she terms the 'AI wilderness', referring to the largely unregulated and pedagogically ambiguous environments in which many learners first encounter AI tools. Her qualitative study identifies significant learner challenges in exercising meaningful autonomy when AI systems are powerful, persuasive, and difficult to critically evaluate. Students in her study frequently adopted AI outputs uncritically, deferring to the authority of the

technology rather than exercising their own judgement. Alm (2024) frames this as a crisis of autonomous choice that educational institutions are not yet adequately equipped to address.

Darvishi et al. (2024) offer perhaps the most analytically rigorous examination of AI's impact on student agency in the current literature. Their large-scale study, published in *Computers & Education*, examines the complex interplay between AI assistance and student agency across multiple dimensions including behavioural, cognitive, and motivational aspects. They find that while AI assistance can enhance certain surface-level indicators of agency, such as task completion rates and resource access, it can simultaneously undermine deeper forms of agency including critical inquiry, intellectual risk-taking, and the development of authentic problem-solving strategies. This distinction between surface agency and deep agency is theoretically important and has implications for how educational outcomes are assessed.

Guarda (2025) addresses the question of epistemic justice alongside autonomy, arguing that generative AI poses specific risks to students' capacity for autonomous knowledge construction. In his analysis, the very efficiency of generative AI tools creates an incentive structure that rewards delegation of intellectual labor to the machine, progressively eroding students' epistemic confidence and their sense of themselves as capable knowledge producers. He further argues that this dynamic has differential effects across social groups, raising concerns about AI's potential to deepen rather than reduce educational inequalities.

Anh et al. (2025) contribute empirical evidence from a case study of high school students in Vietnam, finding that while AI tools were widely used and generally valued by students, their impact on learning autonomy was decidedly mixed. Students reported using AI primarily for task completion and answer-seeking rather than for deeper learning purposes, suggesting that without pedagogical guidance, AI use tends to gravitate toward the path of least intellectual resistance. The authors note that this pattern reflects broader tendencies in adolescent technology use rather than an inevitable feature of AI itself, pointing to the importance of deliberate instructional design.

Shoukat et al. (2026) examine the role of generative AI in enhancing or undermining autonomous learning in higher education and conclude that the balance between these outcomes is highly context-dependent. In their study, students who used generative AI without structured guidance exhibited reduced autonomous engagement over time, while those who used AI within frameworks that required them to critically evaluate, modify, and contextualize AI outputs maintained higher levels of autonomous engagement. These findings suggest that the pedagogical wrapper around AI use is at least as important as the technology itself in determining its effects on autonomy.

Darwin et al. (2024) examine EFL students' perceptions of AI and critical thinking, finding that while students perceived AI as broadly beneficial, they simultaneously expressed concern about its potential to reduce their critical thinking engagement. Students reported a tendency to accept AI-generated language and content without the critical interrogation they would apply to human-produced texts, raising important questions about the long-term implications for students' critical and analytical capacities. The authors argue for explicit critical AI literacy education as a pedagogical response to these tendencies.

Qizi and Fakhridinovna (2026) investigate how AI-powered tutoring systems affect student self-control and learner autonomy, finding that extended use of AI tutors was associated with reduced self-regulatory behavior in some student populations. Students who relied heavily on AI tutoring showed decreased capacity for independent task persistence, suggesting that frequent AI mediation may attenuate the development of self-regulatory muscles that autonomous learning requires.

AI and Student Identity: Transformation and Tension

The literature on AI's effects on student identity is somewhat less voluminous than that on autonomy, but it is rapidly expanding and reveals a rich set of insights about the ways in which AI-mediated learning environments shape, challenge, and transform how students understand themselves.

Thibaud (2026) provides the most comprehensive treatment of AI's impact on student identity formation in the current literature, examining how generative AI tools interact with students' developing sense of academic and personal identity across multiple educational contexts. Drawing on a range of international perspectives, Thibaud identifies three key dynamics: identity extension, in which students expand their perceived capabilities through AI collaboration; identity fragmentation, in which the boundary between the student's own intellectual production and AI-generated content becomes blurred; and identity resistance, in which students actively reject AI tools in order to preserve a sense of authentic intellectual selfhood. These dynamics are not mutually exclusive and may operate simultaneously within the same individual.

Khoudri (2025) focuses on learner identity transformation in digital educational environments, arguing that AI does not merely provide tools for learning but fundamentally reconfigures the relational and epistemic landscape within which learner identity develops. In Khoudri's analysis, the introduction of AI into educational settings shifts the locus of knowledge authority, challenges traditional understandings of student competence, and creates new identity possibilities and pressures. Students must navigate questions of intellectual ownership and authenticity that previous generations did not face in the same form.

Riser (2025) examines generative AI in the high school English classroom and its effects on student identity and teacher practice. His qualitative findings reveal that students using generative AI for writing tasks experienced significant identity negotiations around authorship and creative ownership. For some students, AI use triggered anxiety about whether their work was genuinely their own; for others, it offered a liberating sense of collaborative possibility. Riser concludes that the identity effects of AI in English education are complex and require sensitive pedagogical attention to the meanings students attach to their own writing and creative expression.

Disu (2025) investigates identity and meaning-making among graduate students in creative disciplines who use AI tools, finding that these students engaged in continuous processes of identity work to reconcile their professional and creative self-concepts with AI's increasing presence in their fields. Many experienced what Disu (2025) describes as 'becoming with AI', a form of identity development that neither fully embraces nor rejects the technology but instead constructs a relational mode of creative selfhood that incorporates AI as a collaborator rather than a competitor or substitute. This constructive mode of engagement represents a potentially significant model for how AI-inclusive identity formation might productively unfold.

Li et al. (2025) examine creative personal identity in the age of generative AI through a social-cognitive framework, finding that AI literacy and self-efficacy are significant mediators of the relationship between generative AI use and creative identity development. Students who possessed higher AI literacy and stronger self-efficacy were better able to maintain a coherent and positive creative identity when using AI tools, while those with lower AI literacy and weaker self-efficacy experienced greater identity disruption. The authors highlight the pathway from AI literacy through creative self-efficacy to identity maintenance as a crucial mechanism for educational intervention.

Karimi and Keshvari (2026) study language teachers but their findings have clear relevance to student identity, particularly for mature and mid-career learners. They find that learner autonomy driven by generative AI is a source of emotional vulnerability and identity tension, particularly when AI-generated learning challenges existing professional and personal self-concepts. These tensions, they argue, are not merely cognitive but deeply affective, involving feelings of inadequacy, displacement, and uncertainty that require specific forms of educational and psychological support.

Alfiana and Romadhon (2026) contribute the concept of meta-identity in their examination of curriculum reform in English education. They argue that AI media and assessment practices influence not only students' surface-level academic identities but their meta-identities, defined as their overarching sense of themselves as learners capable of navigating complex, technologically mediated educational environments. When curriculum reforms align AI integration with explicit attention to identity development, students can develop a more resilient and adaptive learner meta-identity. When reforms treat AI as merely instrumental without attending to its identity implications, students may be left to navigate profound identity challenges without adequate support.

Mediating Factors: AI Literacy, Self-Efficacy, Pedagogy, and Context

Across the literature, several mediating and moderating factors consistently emerge as shaping the nature and direction of AI's effects on student autonomy and identity. These factors do not merely moderate statistical effect sizes; they represent the substantive mechanisms through which educational context transforms technological affordance into human developmental outcome.

AI literacy emerges as one of the most consistently cited mediating variables. The conceptual analysis of Mammadova (2026), the empirical study of Li et al. (2025), and the EFL investigation of Darwin et al. (2024) all converge on the finding that students who possess the capacity to critically understand and evaluate AI tools, rather than simply use them, are better positioned to leverage AI for autonomous learning and to maintain coherent identities in AI-mediated environments. AI literacy, in this literature, is not simply a technical skill but a critical and reflective disposition toward AI's capabilities, limitations, and value implications.

Self-efficacy, understood as students' belief in their own capacity to succeed, functions as another crucial mediating variable. Li et al. (2025) demonstrate empirically that self-efficacy mediates the relationship between AI use and creative identity development,

while Huma et al. (2025) find that AI feedback systems that enhance self-efficacy are more likely to promote autonomous learning behaviours. Conversely, Guarda (2025) suggests that AI tools that replace student effort rather than scaffolding it may progressively erode self-efficacy, creating a negative cycle of reduced confidence and increased AI dependence.

Pedagogical design and instructional context are equally prominent in the literature as conditions that shape AI's developmental effects. Ng et al. (2024), Shoukat et al. (2026), and Mulyani and Hermita (2025) all emphasize that the educational outcomes of AI use are not determined by the technology itself but by the pedagogical framework within which it is deployed. When AI is integrated within explicit scaffolding for reflection, critical evaluation, and goal-setting, its autonomy-enhancing potential is maximized. When it is used as a shortcut to task completion without pedagogical structure, it tends to undermine autonomous development.

Institutional and cultural context also shapes AI's educational effects in important ways. Kundu and Bej (2025) highlight how the autonomy-enhancing effects of AI may be particularly significant in the Global South, where students may lack access to other forms of individualized support. Mncube et al. (2026) emphasize that institutional frameworks for AI governance and pedagogical support are critical for realizing AI's positive educational potential in higher education. Anh et al. (2025) and Almohesh (2024) demonstrate that national and cultural contexts, including prevailing norms about the role of students, teachers, and technology in learning, significantly shape how AI tools are used and what their effects are on learner autonomy and identity.

Manglani et al. (2025) and Vieriu and Petrea (2025) both offer broader reviews of AI's educational impact that converge on the importance of contextual factors. Vieriu and Petrea specifically note that AI's effects on students' academic development are mediated by the degree of intentionality with which educational institutions approach AI integration, suggesting that the difference between empowering and disempowering AI environments is primarily a function of deliberate educational choice rather than technological inevitability.

Discussion

The foregoing synthesis of 30 studies reveals a picture of considerable complexity, nuance, and productive tension. AI tools are simultaneously among the most promising and the most challenging developments in contemporary education, and their effects on student autonomy and student identity cannot be adequately characterized by either uncritical enthusiasm or blanket rejection.

The review confirms that AI technologies carry genuine autonomy-enhancing potential, particularly in the domains of personalized learning, adaptive feedback, and the removal of access barriers in under-resourced educational contexts. The meta-analytic evidence of Wang et al. (2026) provides robust quantitative support for these effects, while the qualitative and case study evidence of Ng et al. (2024), Almohesh (2024), Kundu and Bej (2025), and others illuminates the mechanisms through which these effects operate. When students are empowered to use AI as a tool for extending their own cognitive reach, exploring beyond the curriculum, receiving immediate formative feedback, and accessing personalized learning pathways, the technology can meaningfully advance the pedagogical goals associated with autonomous learning.

However, the review also reveals that these benefits are highly conditional and context-dependent. The studies of Darvishi et al. (2024), Alm (2024), Anh et al. (2025), and Shoukat et al. (2026) collectively demonstrate that without deliberate pedagogical scaffolding, critical AI literacy instruction, and institutional support, AI use tends to undermine rather than enhance deep autonomous engagement. The distinction that Darvishi et al. draw between surface agency and deep agency is particularly valuable in this regard: AI may enhance students' capacity to complete tasks, access information, and produce outputs while simultaneously eroding their capacity for independent reasoning, intellectual risk-taking, and genuine self-direction. Educational assessment frameworks that focus on output quality rather than the quality of the learning process may therefore systematically miss the most important effects of AI on student development.

The identity dimension of the analysis reveals perhaps the most profound implications of AI integration in education. The studies of Thibaud (2026), Khoudri (2025), Riser (2025), and Disu (2025) converge on the finding that AI's presence in educational environments raises fundamental questions about intellectual ownership, creative authorship, and authentic selfhood that go to the core of what it means to be a learner. These are not merely philosophical abstractions; they are experienced as lived tensions by students at all educational levels, generating anxiety, self-doubt, and in some cases adaptive resistance that carries its own developmental value.

The concept of 'becoming with AI' articulated by Disu (2025) offers a productive theoretical resource for thinking about identity development in AI-saturated educational environments. Rather than framing the relationship between students and AI as one of competition or displacement, this model suggests that identity development in the AI era involves a complex co-construction in which students develop new understandings of their own capacities, limitations, and creative agency in relational dialogue with AI tools. Whether this co-construction is ultimately empowering or disempowering depends, as the literature consistently suggests, on the conditions under which it occurs, particularly the presence or absence of AI literacy, self-efficacy, and pedagogical support.

Alfiana and Romadhon's (2026) concept of meta-identity provides another productive lens for understanding these dynamics at a curricular level. If AI integration is to support rather than undermine student identity development, educational institutions must attend not only to the technical question of how AI tools are used but to the deeper question of how their use shapes students' evolving sense of themselves as learners. This requires a form of educational intentionality that goes beyond standard AI ethics guidelines to engage with the identity implications of technological integration as a matter of curriculum design.

The mediating role of AI literacy, highlighted across multiple studies, suggests that the binary framing of AI as either beneficial or harmful to student development is analytically insufficient. The educational effects of AI are not properties of the technology but properties of the educational environment in which it is embedded, including the AI literacy capacities of the students who use it. The argument advanced by Mammadova (2026) that AI literacy should be understood as a contemporary form of learner autonomy in itself points toward a reconceptualization of what educational institutions need to cultivate in the AI era: not merely technical AI competence, but the critical, reflective, and evaluative capacities that enable students to remain the authors of their own learning and development even in the presence of powerful AI systems.

Implications

Implications for Educational Practice

The findings of this review carry several important implications for teaching and learning practice. First, educators at all levels need to develop both their own AI literacy and their pedagogical capacity to scaffold students' critical engagement with AI tools. Simply permitting or prohibiting AI use is insufficient; what is required is deliberate instructional design that treats AI literacy as a curricular goal and creates structured opportunities for students to reflect critically on their AI use, its implications for their learning, and its meaning for their developing sense of academic identity.

Second, the literature consistently highlights that AI's autonomy-enhancing potential is realized only when the technology is positioned as a tool for extending student agency rather than replacing it. Educators should design AI-integrated learning tasks that require students to critically evaluate, modify, contextualize, and build upon AI outputs rather than simply accepting them. This requires a pedagogical shift from output-focused to process-focused assessment and a reconceptualization of what counts as productive intellectual work in an AI-augmented educational environment.

Third, the identity implications of AI integration require specific pedagogical attention. Educators in disciplines where creative and intellectual ownership is particularly significant, such as writing, the arts, language learning, and the humanities, should explicitly engage students in reflective activities that address questions of authorship, intellectual ownership, and creative identity. The findings of Riser (2025) about the identity negotiations of high school English students, and the work of Disu (2025) on graduate students in creative disciplines, both suggest that these conversations are already happening informally and that educators can play a meaningful role in facilitating them constructively.

Implications for Curriculum and Policy

At the curricular level, the argument advanced by Alfiana and Romadhon (2026) for explicit attention to meta-identity in curriculum reform has important implications for how AI integration is conceptualized in educational policy. Curriculum designers should ensure that AI integration is accompanied by explicit learning objectives related to AI literacy, critical thinking, and learner identity development, and not only by technical competencies. The risk, highlighted by Guarda (2025) and others, is that AI is incorporated into curricula primarily for its efficiency benefits without adequate attention to its developmental implications.

Mncube et al. (2026) and Vieriu and Petrea (2025) both point to the importance of institutional frameworks for AI governance in higher education. Institutions need policies that go beyond academic integrity guidelines to address the pedagogical,

psychological, and identity implications of AI integration. This includes investment in educator professional development, creation of AI-specific support structures for students, and ongoing monitoring of AI's effects on learner development across multiple dimensions.

The findings of Kundu and Bej (2025) and Anh et al. (2025) suggest that policy attention to equity is essential. AI's effects on student autonomy and identity are not uniform across social and educational contexts; students in the Global South and those in resource-constrained settings may experience both specific benefits and specific vulnerabilities. Educational policies that promote AI integration without attention to equity are unlikely to realize AI's democratizing potential and may inadvertently deepen existing educational inequalities.

Recommendations

Based on the synthesis of the literature, the following recommendations are advanced for educators, curriculum designers, institutional leaders, and policymakers:

First, AI literacy should be embedded as an explicit and assessable learning objective across educational levels and disciplines. This literacy should be understood not only as technical competence but as a critical, reflective, and evaluative capacity oriented toward understanding AI's affordances, limitations, biases, and implications for learner agency and identity.

Second, pedagogical frameworks for AI integration should prioritize the development of deep autonomous engagement. This requires designing tasks that cannot be completed by AI alone, building in structured reflection on AI use and its implications for the quality of learning, and creating assessment approaches that reward intellectual process as well as product.

Third, educational institutions should invest in professional development for educators that addresses not only technical AI competencies but the pedagogical, psychological, and ethical dimensions of AI integration. Educators need both the knowledge to facilitate critical AI literacy and the reflective capacity to address the identity implications of AI in their specific disciplinary and cultural contexts.

Fourth, further research is urgently needed on the longitudinal effects of AI integration on student identity development, on the differential effects of AI use across social, cultural, and disciplinary contexts, and on the specific conditions under which AI tools support rather than undermine deep autonomous engagement. The rapid pace of AI development means that the literature reviewed here, while current, will need continuous updating and critical reassessment.

Fifth, institutional policies on AI should be developed collaboratively with students as well as educators and administrators. The identity negotiations and autonomy challenges documented in the literature are experienced most directly by students, and their perspectives and insights should be central to the development of educational AI governance frameworks.

Conclusion

This narrative literature review has synthesized evidence from 30 scholarly works to construct a comprehensive and critically reflective account of AI's impact on student autonomy and student identity. The review confirms that AI's educational effects are neither uniformly positive nor uniformly negative but are profoundly shaped by the pedagogical, institutional, cultural, and individual contexts in which AI tools are deployed.

The literature documents genuine autonomy-enhancing potential in AI tools, particularly through personalization, adaptive feedback, and the democratization of individualized learning support. At the same time, it reveals significant risks of over-reliance, diminished critical engagement, erosion of epistemic agency, and identity fragmentation when AI is used without adequate pedagogical scaffolding, AI literacy instruction, and institutional support. The distinction between surface agency and deep agency, articulated most sharply by Darvishi et al. (2024), offers an important conceptual resource for distinguishing genuine educational empowerment from mere productivity enhancement.

The identity dimension of the analysis reveals that AI's presence in educational environments raises fundamental and urgent questions about intellectual ownership, creative authorship, and authentic selfhood, questions that are experienced as lived tensions by students and that demand serious and sustained educational attention. The concept of 'becoming with AI' proposed by Disu (2025) and the meta-identity framework developed by Alfiana and Romadhon (2026) both suggest that identity development in the AI era is not simply a matter of individual adaptation but of collective educational reimagination, a challenge that calls upon

educators, institutions, and policymakers to think carefully and creatively about what it means to educate human beings in an increasingly AI-saturated world.

Ultimately, the literature converges on a single overarching insight: the question of whether AI enhances or undermines student autonomy and identity is not a question about technology but about education. The outcomes that students experience are determined not by the capabilities of the technology itself but by the intentionality, wisdom, and ethical commitment with which educators and institutions choose to integrate it into the complex and irreducibly human project of learning.

References

1. Alfiana, N., & Romadhon, M. G. E. (2026). Curriculum Reform in English Education: How Ai Media and Assessment Practices Influence Students' Meta-Identity and Learning Agency. *Journal of Language Sector*, 2(1), 17-39. <https://doi.org/10.69673/jolse.v2i1.223>
2. Alm, A. (2024). Exploring autonomy in the AI wilderness: Learner challenges and choices. *Education sciences*, 14(12), 1369. <https://doi.org/10.3390/educsci14121369>
3. Anh, D. N. H., Chau, N. T. B., & Huyen, N. T. T. (2025). The impact of Artificial Intelligence (AI) on high school students' learning autonomy (a case study of high schools in Thai Nguyen Province). *Journal of Advanced Education and Sciences*, 5(4), 49-58. <https://doi.org/10.64171/JAES.5.4.49-58>
4. Baumeister, R. F., & Leary, M. R. (1997). Writing narrative literature reviews. *Review of general psychology*, 1(3), 311-320. <https://doi.org/10.1037/1089-2680.1.3.311>
5. Boud, D. (2012). *Developing student autonomy in learning*. Routledge.
6. Cinoğlu, H., & Arikan, Y. (2012). Self, identity and identity formation: From the perspectives of three major theories. *Journal of Human Sciences*, 9(2), 1114-1131.
7. Darvishi, A., Khosravi, H., Sadiq, S., Gašević, D., & Siemens, G. (2024). Impact of AI assistance on student agency. *Computers & Education*, 210, 104967. <https://doi.org/10.1016/j.compedu.2023.104967>
8. Darwin, Rusdin, D., Mukminatien, N., Suryati, N., Laksmi, E. D., & Marzuki. (2024). Critical thinking in the AI era: An exploration of EFL students' perceptions, benefits, and limitations. *Cogent Education*, 11(1), 2290342. <https://doi.org/10.1080/2331186X.2023.2290342>
9. Dearden, R. F. (2010). Autonomy and education. In *Education and the Development of Reason (International Library of the Philosophy of Education Volume 8)* (pp. 333-345). Routledge.
10. Desjardins, R. (2015). The Precarious Role of Education in Identity and Value Formation Processes: the shift from state to market forces. *European Journal of Education*, 50(2), 134-146. <https://doi.org/10.1111/ejed.12114>
11. Disu, L. (2025). Becoming With AI: Identity and Meaning-Making Among Graduate Students in Creative Disciplines.
12. Guarda, T. (2025, October). Generative AI in Education: Impacts on Student Autonomy, Critical Thinking, and Epistemic Justice. In *International Conference on Advanced Research in Technologies, Information, Innovation and Sustainability* (pp. 136-147). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-032-16758-3_11
13. Huma, T., Shah, S. S. A., Zaheer, A. M. Z., & Arif, M. (2025). Investigating the Impact of Ai-Driven Feedback Systems on Student Autonomy and Self-Directed Learning. *Social Science Review Archives*, 3(4), 2169-2179. <https://doi.org/10.70670/sra.v3i4.1320>
14. Karimi, M. N., & Keshvari, Z. (2026). Learner autonomy driven by (generative) artificial intelligence as a source of emotional vulnerability and identity tensions in mid-career tertiary-level language teachers. *Teaching and Teacher Education*, 176, 105507. <https://doi.org/10.1016/j.tate.2026.105507>
15. KHOUDRI, I. (2025). Learning with AI: Transformations of Learner Identity in Digital Educational Environments. *Revue LIMES*, 2(2), 183-200. <https://doi.org/10.34874/PRSM.limes-vol2iss2.7402>

16. Kundu, A., & Bej, T. (2025). Empowering students' autonomy in EFL learning: AI innovations in schools of the Global South. *The Electronic Journal of Information Systems in Developing Countries*, 91(6), e70041. <https://doi.org/10.1002/isd2.70041>
17. Li, H., Zhang, Y., Chen, M., Zhao, T., & Jou, M. (2025). Creative personal identity in the age of generative AI: A social-cognitive pathway of AI literacy, self-efficacy, and mindset. *Computers in Human Behavior*, 108838. <https://doi.org/10.1016/j.chb.2025.108838>
18. Mammadova, K. (2026). Redefining student autonomy through AI literacy in the era of generative AI. *Scientific Research International Scientific Journal*, 6(4), 44-47. <https://doi.org/10.36719/2789-6919/57/44-47>
19. Manglani, J., Trivedi, A., & Madhani, J. (2025). The impact of artificial intelligence (AI) on education—A review paper. *Digital Transformation and Sustainability of Business*, 570-573.
20. Mncube, D. W., Maphalala, M. C., & Mkhasibe, R. G. (2026). Artificial intelligence in higher education: Supporting self-directed learning and student autonomy. *Turkish Online Journal of Distance Education*, 27(1), 275-290. <https://doi.org/10.17718/tojde.1668721>
21. Mulyani, E. A., & Hermita, N. (2025). AI-BASED ADAPTIVE LEARNING MODELS: THEIR INFLUENCE ON LEARNING PERSONALIZATION AND STUDENT AUTONOMY. *Al-Hijr: Journal of Adulearn World*, 4(3). <https://doi.org/10.55849/alhijr.v4i3.1064>
22. Ng, D. T. K., Tan, C. W., & Leung, J. K. L. (2024). Empowering student self-regulated learning and science education through ChatGPT: A pioneering pilot study. *British Journal of Educational Technology*, 55, 1328-1353. <https://doi.org/10.1111/bjet.13454>
23. Ouchene, N., & Belmekki, A. (2025). Motivation, Autonomy, and AI: Exploring Student's Attitudes of Intelligent Tools in the Learning Process. *Passerelle*, 14(1), 396-416.
24. Qizi, H. A. U., & Fakhridinovna, H. L. (2026). INVESTIGATING HOW STUDENT SELF-CONTROL AND LEARNER AUTONOMY ARE AFFECTED BY AI-POWERED TUTORING SYSTEMS. *Инновационные методики в обучении иностранным языкам: глобальные тенденции и опыт университетов шёлкового пути*, 1, 1585-1590. <https://doi.org/10.24412/cl-37493-2026-1-1585-1590>
25. Rashed Ibraheam Almohesh, A. (2024). AI application (ChatGPT) and Saudi Arabian primary school students' autonomy in online classes: Exploring students and teachers' perceptions. *International Review of Research in Open and Distributed Learning*, 25(3), 1-18. <https://doi.org/10.19173/irrodl.v25i3.7641>
26. Riser, P. (2025). Generative artificial intelligence (AI) in the high school English classroom and its effect on students' identity and teachers' practice. *English in Education*, 59(2), 176-192. <https://doi.org/10.1080/04250494.2025.2491361>
27. Shoukat, R., Parveen, K., Iqbal, A., Shafiq, M., & Qibing Sr, H. (2026). The Role of Generative AI in Enhancing or Undermining Autonomous Learning in Higher Education. *Journal of Educators Online*, 23(1), n1.
28. Thibaud, E. (2026). The Impact of Generative AI on Student Identity Formation. In *International Perspectives of Generative AI in Education: Insights from Research in Diverse Educational Contexts* (pp. 327-344). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-99804-1_19
29. Torraco, R. J. (2005). Writing integrative literature reviews: Guidelines and examples. *Human resource development review*, 4(3), 356-367. <https://doi.org/10.1177/1534484305278283>
30. Vieriu, A. M., & Petrea, G. (2025). The impact of artificial intelligence (AI) on students' academic development. *Education Sciences*, 15(3), 343. <https://doi.org/10.3390/educsci15030343>
31. Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. *Expert systems with applications*, 252, 124167. <https://doi.org/10.1016/j.eswa.2024.124167>

32. Wang, W., Wang, Y., Chen, J., Wang, X., Zhang, H., Guo, C., & Peng, Y. (2026). The effectiveness of AI-supported personalized feedback on students' learning outcomes and motivation: a meta-analysis. *Journal of Educational Computing Research*, 64(3), 724-755. <https://doi.org/10.1177/07356331251410020>



2026 by the authors; Journal of Global Social Transformation (JGST). This is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).