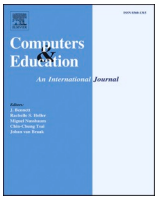
Computers & Education 253 (2026) 105687



Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu





Redefining and measuring student agency in AI-assisted learning: Development and validation of the agentic engagement with AI (AE-AI) scale

Yun Dai [a,*], Suya Liu [a], Sihan Zhou [a], Sichen Lai [a], Ang Liu [b], Cher Ping Lim [c]

[a] Department of Curriculum and Instruction, The Chinese University of Hong Kong, Shatin, Hong Kong Special Administrative Region
[b] School of Mechanical and Manufacturing Engineering, University of New South Wales, Australia
[c] Department of Curriculum and Instruction, The Education University of Hong Kong, Shatin, Hong Kong Special Administrative Region

ARTICLE INFO

Keywords:
Learner agency
Agentic engagement
Human–AI co-agency
Generative AI
Epistemic responsibility
adaptability
Scale validation

ABSTRACT

Student agency has become a central concern in learning with artificial intelligence (AI). With growing capability and adaptiveness, Generative AI (GenAI) tools are no longer merely passive tools but responsive participants in learning. This shift makes agency relational and reciprocal: students exercise agency through intentional engagement with AI, while AI responses shape how students express, adjust, and develop that agency. Existing instruments, however, continue to frame agency as student-centered control over external tools and environments, positing students as the primary initiators and regulators of learning. This misalignment leaves the relational and reciprocal nature of agency in AI-assisted learning unmeasured. To address this gap, the present study adopts the construct of agentic engagement and conceptualizes student agency in AI-assisted learning as a *proactive*, *adaptable*, and *reciprocal* process enacted through ongoing exchanges between human intentionality and machine adaptivity. To develop the Agentic Engagement with AI (AE-AI) Scale, an integrated inductive–deductive approach was used to generate an initial pool of 28 items from student interviews ($N = 26$). The scale was refined and validated through expert review, exploratory factor analysis ($N = 340$), confirmatory factor analysis ($N = 256$), and criterion validity checking. The final 16-item scale comprises four factors: *Adaptive Direction*, *Critical Integration*, *Cross-Source Inquiry*, and *Reflective Calibration*. Together, these factors redefine student agency as a process of human–AI *co-agency* grounded in adaptability, epistemic responsibility, and distributed inquiry. The study provides a validated instrument for observing student agency and opens new directions for examining the evolving synthesis of human and machine intelligence in education.

1. Introduction

The rapid adoption of AI tools among university students has brought learner agency to the forefront of education research and practice (Dai et al., 2023; Darvishi et al., 2024). Student agency generally refers to the capability to initiate and lead one's learning experiences, while influencing his/her environments (Scott, 2008). It is arguably believed that the extent to which students can exercise agency decides the depth of their learning, the authenticity of their engagement, and their ability to navigate challenges in

* Corresponding author.
E-mail address: yundai@cuhk.edu.hk (Y. Dai).

https://doi.org/10.1016/j.compedu.2026.105687
Received 8 October 2025; Received in revised form 17 May 2026; Accepted 6 June 2026
Available online 11 June 2026
0360-1315/© 2026 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

AI-assisted environments (Guo et al., 2025; Uanachain & Aouad, 2025). Given its significance, educators and researchers have proposed various conceptualizations of student agency, including postdigital learner agency (Code, 2025), epistemic agency (Wu et al., 2025), and the student-AI agency framework (Vincoli et al., 2025). A concurrent trend is to develop new instruments to operationalize and measure student agency, such as the Student Learning Agency in GenAI-supported Contexts (SLA-GAI) Scale by Xia et al. (2025) and a learning-approach-centered analytical framework by Yang et al. (2024).

Nevertheless, a substantial gap persists between theoretical advances and existing instruments of student agency in AI-assisted learning. As GenAI tools can generate sophisticated outputs and respond adaptively to human inputs, recent scholarship has increasingly recognized the notion of co-agency between humans and AI (Nordström et al., 2023; Satyanarayan & Jones, 2024). In this view, human–AI interactions are *reciprocal* and *transactional*, with each party shaping the other's actions throughout the learning process (Code, 2025). Student agency, then, is not reflected only in intentional efforts to lead interactions with AI. It is also reflected in students' *adaptability*: reading and assessing AI outputs, adjusting interactional strategies, and refine AI-assisted learning experiences over time. Most existing scales, however, continue to conceptualize agency as student-centered control, emphasizing a unidirectional flow of agency from students to external tools and environments (e.g., Guo et al., 2025). This framing still rests on the assumption that the learning environment is a static structure to be regulated, rather than a responsive system that participates in the interaction. New instruments are therefore needed to capture how students co-construct learning through intentional, adaptive, and reciprocal engagement with AI.

This study develops and validates the Agentic Engagement with AI (AE-AI) Scale, an instrument designed to measure students' proactive efforts in leading and adapting their learning with AI. The construct is grounded in agentic engagement theory (Reeve et al., 2020; Reeve & Tseng, 2011), which views students as active participants who shape their learning conditions through proactive and transactional participation, rather than passive recipients of instruction. Informed by this theoretical lens, the construct of AE-AI captures the ways students direct, negotiate, evaluate, and reflect on their interactions with AI tools in an adaptable and reciprocal manner. The scale development followed a rigorous process that included a literature review, item generation, expert review, pilot testing, and psychometric validation. The resulting AE-AI Scale offers researchers and educators a theoretically grounded, empirically validated tool for examining agency in AI-assisted learning. The study also offers important implications for conceptualizing and understanding human–AI collaboration in the evolving technological landscape.

2. Literature review

2.1. Student agency in AI-assisted learning

The use of AI to assist learning has evolved through several phases since the 1970s, beginning with rule-based Intelligent Tutoring Systems and later expanding into adaptive systems and learning analytics for data-driven personalization (Holmes & Tuomi, 2022; Nwana, 1990; Zhai et al., 2021). Since the early 2020s, generative AI (GenAI) has introduced a new paradigm. Powered by large language models (LLMs) and other technological advances, GenAI can generate text, images, code, and other multimodal content with considerable sophistication (George, 2023).

Student agency has long been a key topic in the pursuit of more intelligent and autonomous education systems, especially with the advent of GenAI technologies (Dai et al., 2023; Darvishi et al., 2024; Mouta et al., 2025). Educators and developers have long explored how to leverage the affordance of AI technologies, such as personalization, adaptivity, and automation, to enhance student engagement, autonomy, and sense of ownership in the learning process (Brod et al., 2023; Marín et al., 2024). Meanwhile, scholars have cautioned that complete automation of learning risks diminishing students' capacity to make choices and exercise judgement (James et al., 2022). Heavy use of technological support can shift control away from students, as systems make decisions about content, pace, and direction that students might otherwise make themselves (Molenaar, 2022).

The unique affordances and limitations of GenAI tools add an extra layer of complexity for student agency, challenging the conventional conceptualization of agency. First, GenAI's sophisticated capability expands what AI can offer students. Unlike earlier systems that primarily analyzed data or recommended resources, GenAI can produce tailored explanations, feedback, examples, and drafts on demand (McGrath et al., 2024). This opens more diverse ways for students to seek help and shape their learning paths, supporting tasks ranging from tutoring, coding, document processing, and others (Dai et al., 2023). At the same time, current GenAI models may produce inaccurate, biased, de-contextualized, or overly general outputs (Rudolph et al., 2023). These limitations require students to critically assess what AI provides and to use it responsibly. Without such judgement, students risk overreliance on AI or passive consumption of its outputs, which can erode their critical thinking, originality, and independent skills (Bontcheva et al., 2024). In this sense, student agency becomes not only a matter of maximizing the affordances of GenAI but also of safeguarding learning integrity against its limitations (Nopas, 2025).

Second, GenAI's adaptiveness reshapes how students interact with AI in the learning process. Unlike earlier AI applications that responded to pre-defined inputs or fixed student data, GenAI adjusts its responses in real time to students' open-ended prompts, follow-up questions, feedback, and data (Khine, 2024). Students now engage with GenAI in many forms, from conversational exchanges to building their own AI agents and integrating GenAI into personalized workflows (Lai et al., 2025). Across these forms, the quality of AI outputs depends heavily on what students bring to the interaction (Dai et al., 2023; Mahdavi Goloujeh et al., 2024). Thoughtful prompts and applications lead to richer outputs, while vague or passive inputs lead to weaker ones. The interaction is thus reciprocal: each input from the student shapes the AI's response, and each AI response in turn shapes the student's next move (Dai et al., 2023). Through this back-and-forth, students and GenAI tools mutually shape the flow and direction of learning in an ongoing, transactional exchange.

These two features of GenAI, sophisticated capability and adaptiveness, together reshape what AI contributes to learning. GenAI no longer functions only as a passive tool that executes commands, but as a responsible participant that contributes to, redirects, and shapes the learning process. This changing role has been recognized in recent arguments about agentic AI (Kostopoulos et al., 2025), "intelligence as agency" (Satyanarayan & Jones, 2024), and "agency without intelligence" (Floridi, 2025). In response, scholars have introduced concepts such as student-AI co-agency (Jiang et al., 2024; Roe & Perkins, 2024), shared agency between students and algorithmic agents (Code, 2025), and joint attention between humans and AI (Anton et al., 2025; Cosentino et al., 2025). Although these concepts differ in emphasis, they share a common view: student agency in AI-assisted learning is a *relational*, *reciprocal*, and *transactional*. It emerges from the ongoing exchange between students and AI, rather than from students' actions alone. This view challenges the traditional *unidirectional* account of student agency, which positions students as the primary initiators and regulators of their learning environments (Mairitsch et al., 2023). Such accounts emphasizes individual autonomy and control but overlooks how intelligent systems act, respond, and shape learning trajectories (Roe & Perkins, 2024). To capture this complexity, theoretical frameworks must account for both students' proactive efforts and the reciprocal dynamics of human–AI interaction (Holstein et al., 2019; Vincoli et al., 2025).

2.2. Conceptualizing student agency via the engagement lens

Among existing frameworks of student agency, *agentic engagement* theory offers a strongest conceptual fit for the current study (Reeve, 2013; Reeve et al., 2020; Reeve & Tseng, 2011). Grounded in motivational theories, this theory argues that students are motivated agents who do not merely respond to instruction but actively shape their learning conditions (Reeve, 2013). It suggests the dual characters of agency, that is, proactive on the student's side and transactional in the student's interactions with the learning environment, which is well aligned with the co-agency in AI-assisted learning. Besides, it contextualizes agency within the broader framework of *student engagement*, which provides a well-established foundation for operationalizing and measuring agency in concrete, observable actions (Trowler, 2010).

Student engagement refers to learners' active involvement and emotional investment in academic activities (Finn & Zimmer, 2012). Conventionally, it has been viewed as a meta-construct comprising three dimensions: affective, behavioral, and cognitive engagement, which respectively concern students' emotional connection, actions, and mental processes in learning (Fredricks et al., 2004). Yet, this tripartite framework has been criticized for portraying students primarily as *reactors* who comply with or respond to assigned tasks (Klemenčič, 2017). It overlooks students' constructive and initiatory roles as co-creators of educational experiences, who actively influence and personalizing their learning environments to align with their goals and values (Pineda-Báez et al., 2019). Empirical evidence from motivational research supports this critique: students do not merely respond to educational structures but can exercise agency by expressing needs, asking questions, providing feedback, and directing parts of their learning journey (Bandura, 2014; Klemenčič, 2020; Koenigs et al., 1977; Reeve & Tseng, 2011; Shogren et al., 2015).

To address this gap, Reeve (2012) proposes a fourth dimension of student engagement, "agentic engagement," defined as "the process in which students proactively try to create, enhance, and personalize the conditions and circumstances under which they learn" (p.161). This dimension is distinguished by two features. First, it is proactive: agentically engaged students act in anticipation of teacher-led activities and contribute actively to how their learning unfolds (Reeve & Tseng, 2011). This proactive process is also "an ongoing series of dialectical transactions between students and teachers" (Reeve, 2013, p. 580). Second, it is transactional: learning becomes an ongoing exchange in which students and teachers influence each other through reciprocal causation, where each party's behaviors influence the others' (Cohen et al., 2020; Reeve & Shin, 2020). By continually influencing their learning environments, students maintain authorship and ownership of their learning, thereby exercising agency.

These two features of agentic engagement, proactive contribution and transactional exchange, align with what AI-assisted learning demands. As GenAI tools respond to human inputs in adaptive ways, they transform learning environments into interactive spaces where student inputs and AI outputs constantly (re)shape each other (McGrath et al., 2024). GenAI thus functions as a new kind of "others" in the learning environment, a non-human participant that engages with students reciprocally (Brod et al., 2023). Student-AI interaction becomes *bidirectional* and *co-constructive*: students initiate, direct, and evaluate AI responses, while AI outputs prompt further reflection and adjustment from students (Guo et al., 2025). Agency in this context is therefore not only proactive but also adaptive, as students must respond to AI outputs, adjust their strategies, and steer the exchange towards their learning goals. As such, the framework of agentic engagement provides a strong lens for observing and operationalizing these expressions of agency, which form the basis for the scale developed in this study.

2.3. Existing scales of student agency in technology-enhanced learning

The literature reveals two main approaches to measuring student agency in technology-mediated learning: inferring agency from proxy constructs and developing direct measurement scales. The first approach draws on established constructs from outside AI contexts, such as learner autonomy (Zarrinabadi et al., 2021), self-regulated or self-directed learning (Gupta et al., 2024; Marín et al., 2024), or metacognitive regulation (Wise et al., 2016). Related instruments, such as the Agency of University Students (AUS) scale (Jääskelä et al., 2017) and the Agency for Learning scale (Code, 2020) have also been used in studies of AI-assisted learning, although they were developed for general university contexts. These instruments capture students' active participation, self-regulation, and task completion, treating AI as part of the broader learning environment rather than a distinct interactional partner. As a result, they may not capture the context-specific strategies through which students enact agency when working with AI tools (Xia et al., 2025).

A recent trend is the development of context-specific scales for AI-assisted learning. The Student Learning Agency in GenAI-

supported Contexts (SLA-GAI) Scale is a comprehensive 34-item instrument organized around ten factors, including self-cognition, goal setting, self-adjustment, selective action, self-efficacy, and volition (Xia et al., 2025). Its breadth makes it valuable for capturing multiple facets of agency, although its factor structure appears to lack a unifying framework across the ten factors. Guo et al. (2025) take a different approach with a brief three-item measure (initiative, perceived ownership, and mental effort). While useful for effective administration, this instrument captures only a narrow slice of agency and would benefit from extension to cover the fuller range of agentic expressions. Another effort, the Critical Reflection and Agency in Computing Index (Padiyath et al., 2025), focuses on ethical and sociotechnical reflection among undergraduate computing students. This instrument leaves open the question of how agency is exercised and manifested by students across disciplines when learning with AI.

Taken together, existing scales have made important contributions, but a gap remains. None is grounded in a framework that captures both students' proactive efforts and the reciprocal, transactional dynamics of human–AI interaction. Against this background, this study, by adopting the framework of agentic engagement, seeks to develop a theoretically grounded and context-specific instrument for measuring student agency in AI-assisted learning.

3. Methods

The present study aims to develop and validate the Agentic Engagement with AI (AE-AI) Scale in AI-assisted learning environments. The scale development included three phases: (1) generation of an initial item pool, (2) expert review, and (3) pilot testing. Then, the instrument was psychometrically validated through factor analyses, and its criterion validity was assessed by examining the correlation with AI literacy.

3.1. Participants

Participants in the study included 26 interviewees whose responses generated the questionnaire item pool, and a total of 596 questionnaire respondents for exploratory factor analysis (EFA, $N = 340$), confirmatory factor analysis (CFA, $N = 256$), and a criterion validity test. Data were collected at two universities in China. All participants provided informed consent prior to participating in the study.

Using a purposive sampling strategy, the research team collaborated with two university liaisons to recruit students for the study. Recruitment information was distributed via social media popular among students (i.e., WeChat and online forums). Among 34 students who volunteered, participants who reported regular use of GenAI tools (e.g., Doubao, DeepSeek, ChatGPT) were identified. Then, using a maximum variation sampling strategy, 26 students were selected to represent different academic disciplines, genders, and year levels to capture variations in student responses and enhance the transferability of the findings. The participant received a compensation of RMB 200 (≈ USD 29) for a 1.5 – 2-h interview.

With assistance from university administrative staff, we sent cold emails through the organizational email systems to recruit questionnaire participants. To ensure broad coverage of participants, we purposefully approached students from different majors in different year levels. The email outlined the purpose and scope of the study, and students provided electronic informed consent before participation. Eligibility was restricted to students with prior experience of using GenAI tools in their learning, to ensure that responses reflected authentic engagement. Questionnaire responses were collected via the online survey platform Wenjuanxing in China. All responses were collected anonymously to protect participants' privacy. Participation was incentivized with a small e-voucher (RMB 10; ≈ USD 1.3) upon completion. Across the recruitment period, 644 submissions were received from 1608 recruitment emails; the response rate was approximately 40%, which is acceptable for online educational survey research (Wu et al., 2022). After data screening and cleaning, 596 valid cases remained for analysis.

To ensure independence between EFA and CFA, the cleaned dataset was randomly split into two non-overlapping cohorts for EFA ($N = 340$) and CFA ($N = 256$) test. Following Singh et al. (2016)'s suggestion, each cohort sampled over 200 students and an item-to-participant ratio of at least 1:10 to ensure sample adequacy for factor analysis. Table 1 summarizes the demographic characteristics of participants.

Table 1
Demographic characteristics of the participants.

	Interviewees ($N = 26$)	EFA ($N = 340$)	CFA ($N = 256$)
Gender			
Female	13	239	169
Male	13	101	87
Major of study			
STEM	13	129	97
Humanities and social sciences	13	211	159
Year Level			
Year 1	3	56	38
Year 2	5	113	81
Year 3	6	88	77
Year 4	6	83	60

3.2. Instruments

The questionnaire comprised three parts. It began with two screening questions to confirm eligibility ("What AI tools have you used for academic tasks?" and "What is your average time of using AI per week?"). Only participants reporting prior academic use of AI tools were able to proceed to the main survey. The questionnaire then included the two scales administered in this study. The first questionnaire was the instrument under development and validation: the AE-AI Scale.

The other questionnaire was the AI Literacy Scale (AILS) developed by Wang et al. (2022) to examine the criterion validity of AE-AI Scale. AI literacy refers to the competence that enables people to use, interact with, and critically evaluate AI technologies (Long & Magerko, 2020). Literature shows that AI literacy and student agency are conceptually related but distinct constructs (Avsec & Rupnik, 2025). Literacy is a competence that supports engagement. Students with higher AI literacy (e.g., better understandings of AI's strengths and limitations) tend to engage more meaningfully with AI tools (Ng et al., 2021). Agentic engagement reflects strategic behaviors that often accompany sophisticated knowledge and skills (Guo et al., 2025; Xia et al., 2025). Given these theoretical connections, we hypothesize a significant correlation between AI literacy and agentic engagement, through which AILS measurement can establish criterion validity for the AE-AI Scale. The AILS included 12 items across four dimensions of awareness, usage, evaluation, and ethics. This scale has been validated among adult learners (Wang et al., 2022). The scale was checked for reliability with the participants of this study, with a Cronbach's α of .826.

3.3. Scale development

Following established scale-development guidance (Boateng et al., 2018; Zhou et al., 2024), we used an integrated inductive-deductive procedure to generate the initial item pool. The inductive phase began with stimulated recall interviews with 26 students (Leighton, 2017). Before the interview, each participant submitted ten conversation threads with GenAI tools used for academic tasks, with each thread containing at least five back-and-forth turns. Using these chat logs as interview stimuli, participants were guided to recall specific considerations, actions, and decision points when they planned prompts, refined instructions, evaluated outputs, and adjusted strategies during their interaction with GenAI. Interviews were audio-recorded and transcribed verbatim. The transcripts were coded to identify recurring action statements. Two coders independently conducted open coding on an initial subset of transcripts (eight transcripts) to develop a preliminary codebook, which was then refined through discussion and applied to the full dataset. The interrater reliability was acceptable (Cohen's $\kappa = .72$). Discrepancies were resolved through discussion to reach consensus. Throughout coding, we monitored conceptual saturation and observed that no substantively new action codes emerged in the later transcripts.

The identified action codes were used to draft items that captured concrete, observable strategies rather than abstract dispositions. To enhance stability, only action statements reported by at least two participants were translated into behavior-focused items (e.g., "I modify AI outputs to fit the specific contexts of my tasks"). Deductively, we then reviewed agentic engagement theory and relevant instruments in technology-mediated learning to refine item wording, check construct coverage, and add items for underrepresented behaviors (Kim & Song, 2023; Mameli & Passini, 2019; Reeve, 2013; Dai & Lai, 2025). When inductive and deductive items overlapped conceptually, we merged them and retained the clearer wording anchored in students' language. The consolidation yielded the 28-item pool. Appendix A reports the primary origin of each item.

A 6-point Likert scale was adopted for item responses, ranging from (1) strongly disagree to (6) strongly agree. This format was chosen to capture meaningful variation in students' levels of agreement while avoiding excessive response options that might burden participants (Dörnyei & Dewaele, 2022). Following the guidance of Chomeya (2010), the omission of a neutral midpoint was intended to reduce the tendency for respondents to select a 'safe' middle option, thereby encouraging more definitive and informative responses. A sample item of the 6-point Likert scale is provided in Appendix B.

The initial 28 items were reviewed by four experts with at least five years of research experience in student agency and AI-assisted learning to assess content validity. Each expert independently rated item relevance on a four-point Likert scale (1 = Not relevant, 4 = Highly relevant) and provided qualitative feedback on wording clarity and theoretical alignment. We computed the Content Validity Index (CVI) to summarize expert agreement, adopting a cut-off of 1.00 for acceptable content validity (Polit & Beck, 2006). Four items (#5, #15, #24, #27) were identified as theoretically irrelevant and removed. Based on experts' qualitative feedback, five items (#4, #10, #16, #20, #26) were reworded to improve clarity and interpretability. This review resulted in a refined pool of 24 items for pilot testing and psychometric validation. Appendix A documents the item refinement process.

Before piloting, the entire questionnaire, including AE-AI and AILS scales, was translated into Chinese and qualitatively piloted with five students to refine the wording. The translation was conducted by a Chinese-English bilingual researcher with expertise in student agency and AI literacy to ensure accurate meaning. It was then back-translated by another Chinese-English bilingual researcher trained in translation and linguistics, to enhance semantic equivalence.

3.4. Questionnaire validation

After developing and piloting the questionnaire, it was subjected to psychometric validation via EFA and CFA, as well as a criterion validity test. The questionnaire data were analyzed using Mplus 8 and RStudio. Data was screened for missing values and outliers and checked for normality prior to statistical analysis.

EFA was conducted with 340 questionnaire responses using Maximum Likelihood (ML) extraction and direct oblimin rotation to explore the latent structure of the scale (Singh et al., 2016). The factorability of data was examined through the Kaiser-Meyer-Olkin

(KMO) measure of sampling adequacy and Bartlett's Test of Sphericity. Higher KMO values indicate stronger adequacy and a significant Bartlett's test support factorability (Saini & Abraham, 2021; Zhou et al., 2024). To determine the number of latent factors, parallel analysis and the scree plot were used. Items were removed if they met the following criteria: communality <.30, primary factor loadings < |.40|, cross-loadings >.30, and a difference of < .15 between competing loadings (Hou et al., 2025; Schmitt & Sass, 2011). Items failing these criteria were removed step by step, and the analysis was repeated until a clear and interpretable factor structure was obtained.

Subsequently, CFA was performed with a new dataset of 256 responses to examine whether the data fit the hypothesized number of factors from EFA and the structure of factor loadings. Before analysis, assumptions for CFA were checked, including multivariable normality and sufficient sample size ($N > 200$). Multivariate normality was assessed using Mardia's test to select the appropriate estimation method (Kline, 2023). Model fit was used to determine whether the data fit the model well. The criteria of fit indices included: the Chi-Square/degree of freedom (χ^2/df) ≤ 3, the Comparative Fit Index (CFI) $\geq .90$, Tucker Lewis Index (TLI) $\geq .90$, the standardized root mean square residual (SRMR) $\leq .08$, and the root mean square error of approximation (RMSEA) $\leq .08$ (Kline, 2023). To improve the model fit, modification indices were considered only when (1) the pair of items loaded on the same factor, (2) they conceptually shared residual variance due to semantic overlap, and (3) the modification index was substantially large (Rodríguez-Castro et al., 2017; Saini & Abraham, 2021).

After CFA, the reliability of the AE-AI Scale was assessed using Cronbach's α for the total scale and each subscale, as well as composite reliability (CR) for each latent factor. Values above .70 for both indices indicated acceptable reliability (Hair et al., 2009).

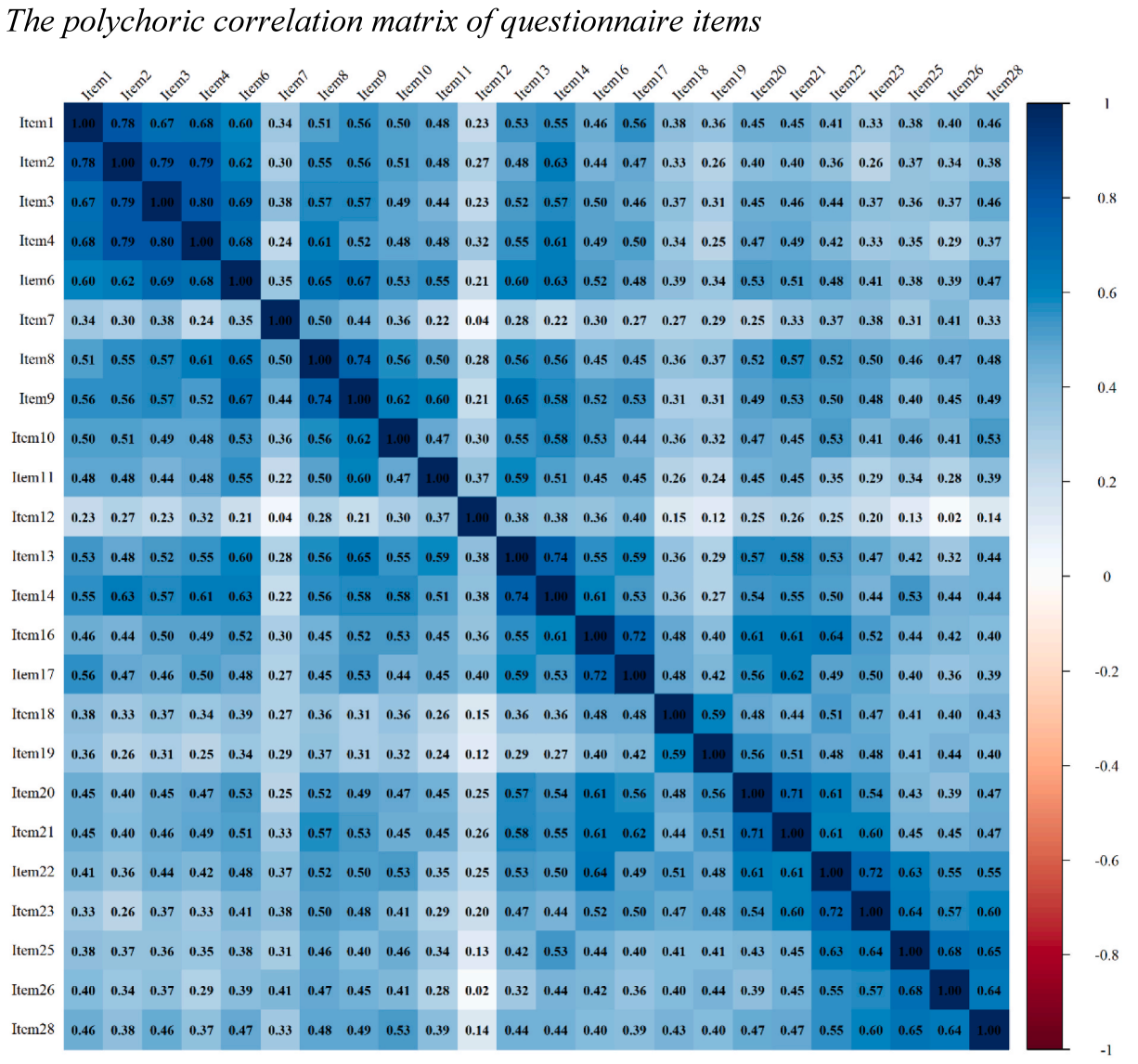


Fig. 1. The polychoric correlation matrix of questionnaire items.

Criterion validity was examined by calculating the Pearson correlation between AE-AI and AILS scores, with coefficients of .10, .30, and .50 interpreted as small, medium, and large effect, respectively (Cohen, 1988). All significance tests were two-tailed with an alpha level of .05. Discriminant validity between the AE-AI factors and the AILS factors was examined using the heterotrait–monotrait ratio of correlations (HTMT) by contrasting heterotrait item correlations (between constructs) with monotrait item correlations (within constructs). HTMT values below .85 were taken to indicate satisfactory discriminant validity (Henseler et al., 2015).

4. Results

4.1. Preliminary analysis

Before proceeding with EFA, the dataset was screened for missing values, outliers, and normality. Listwise deletion was applied to handle missing data, resulting in the removal of 29 cases (Cox et al., 2014). Outliers were then identified using Z-scores, with cases exceeding |3.29| on any item treated as potential outliers and removed (Tabachnick & Fidell, 2013). This resulted in the exclusion of an additional 19 cases, yielding a final sample of 596 valid responses for analysis. Descriptive statistics were calculated for all items. Skewness values ranged from −.74 to −.17, and kurtosis values ranged from −.72 to 1.30. All skewness values fell within ±2, and kurtosis values were between ±7, therefore indicating no substantial deviation from normality (Kim, 2013).

4.2. Exploring the factor structure

To examine the latent structure of the scale, EFA was performed with a sample of 340 participants. Polychoric correlation was used to suit the ordinal nature of the Likert scale (Holgado-Tello et al., 2008). Fig. 1 illustrates the correlation matrix among the 24 items. Two flagged questionnaire items (Items 7 and 12) were weakly correlated with all other items; these were removed prior to factor extraction. The suitability of the data for EFA was then evaluated. The KMO measure of sampling adequacy was .93, and Bartlett's test of sphericity was significant ($\chi^2 = 4799.52$, $p < .001$), indicating that the data were appropriate for factor analysis. EFA was conducted using the Maximum Likelihood extraction method with direct oblimin rotation, as AI agentic engagement was expected to be a multidimensional construct with correlated factors. A parallel analysis was also performed in R Studio using the *psych* package (Horn, 1965; Iacobucci et al., 2022). The parallel analysis suggested a four-factor solution, with the first four observed eigenvalues exceeding those derived from randomly generated data (see Fig. 2). Together, these four factors accounted for 64.33% of the total variance in the initial solution.

Alternative three- and five-factor solutions were also examined. The three-factor solution yielded conceptually overlapping factors that failed to maintain discriminant content. In contrast, the five-factor solution yielded an underpopulated fifth factor, with fewer than three items showing salient loading, which limits its interpretability (Costello & Osborne, 2005). The four-factor solution was therefore retained because it offered the strongest balance of theoretical coherence, empirical interpretability, and factorial parsimony.

Problematic items were then removed stepwise to obtain a clean factor solution, as documented in Appendix A. In total, six additional items were deleted across multiple rounds of EFA: four items (Items 6, 8, 10, and 17) because of weak primary loadings (<|.40|), and two items (Items 16 and 22) because of cross-loadings greater than .30 or insufficient differences between competing loads (<.15). After each round of deletion, the analysis was re-run to confirm the stability of the emerging factor structure. The final solution yielded a four-factor model with 16 items (see Table 2). These factors accounted for 66.54% of the total variance.

Factor labelling was then conducted by examining the content of the items loading on each factor. **Factor 1** was labelled *Adaptive*

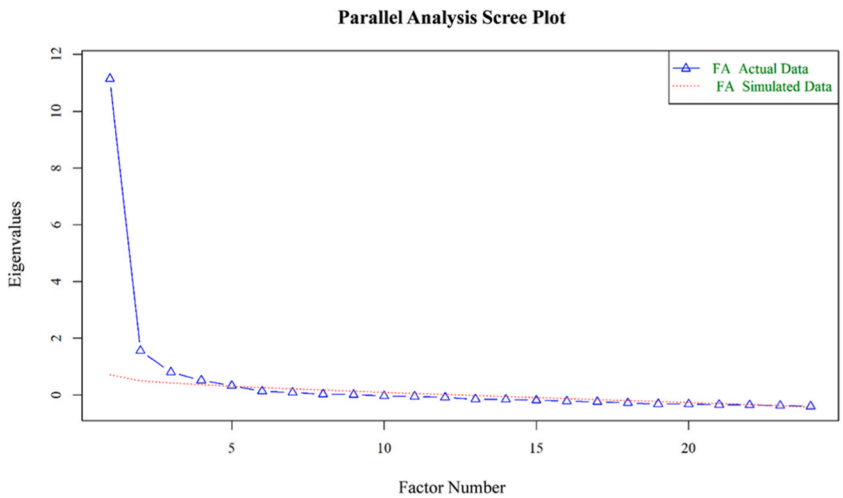


Fig. 2. The scree plot of parallel analysis.

Direction because the items describe students' active efforts to direct AI and adapt their instructions to align AI outputs with their learning goals. **Factor 2** was named *Critical Integration*, as the items capture students' deliberate evaluation, selection, and modification of AI outputs to fit their own knowledge, judgement, and task contexts. **Factor 3** was labelled *Cross-Source Inquiry*, as the items reflect efforts to consult non-AI sources, including peers, mentors, experts, and other materials, to verify, supplement, or deepen AI-assisted work. **Factor 4** was identified as *Reflective Calibration*, as the items describe students' ongoing reflection on whether AI use serves their learning goals and their adjustment of future AI use based on this reflection. Together, these four factors represent distinct yet interrelated dimensions of students' agentic engagement with AI.

Given the use of oblique rotation, interfactor correlations were examined. As shown in Table 3, the factors were moderately correlated ($r = .40–.59$), with the highest correlation observed between Critical Integration (CI) and Reflective Calibration (RC) ($r = .59$). All values were well below the .80-.85 range commonly considered indicative of poor discriminant validity (Brown, 2015; Watkins, 2018). These moderate correlations indicate that the dimensions are related yet empirically distinct, supporting the specification of correlated latent factors in the subsequent CFA.

4.3. Confirmatory factor analysis

To confirm the latent factor structure, CFA was performed with a sample of 256 participants. Maximum likelihood (MLR) was applied as a robust estimation method to address potential violations of multivariate normality (Mardia's skewness and kurtosis, $p < .001$) (Antonietti et al., 2023). The initial four-factor model indicated the need for a theoretically justified modification that allowed the residuals of AD1 and AD2 to correlate, given their semantic similarity in measuring prompt refinement. The revised model showed a good fit ($\chi^2/df = 2.54$, $p < .001$, CFI = .927, TLI = .909, RMSEA = .077, 90% CI [.066, .090], SRMR = .073). Fig. 3 illustrates the model with specifications and standardized parameters. All standardized factor loadings were significant and exceeded .5, meeting recommended thresholds (Hair et al., 2009). The finalized version of the AE-AI Scale is presented in Appendix C.

4.4. Reliability analysis

Reliability analyses indicated that the AE-AI Scale demonstrated strong internal consistency, with a Cronbach's α of .89 for the overall scale. Subscale reliabilities were also acceptable, ranging from .78 to .87, as shown in Table 4. To complement Cronbach's α, composite reliability (CR) values were calculated for each latent factor based on the CFA results. The CR values were .85 for Adaptive Direction (AD), .84 for Critical Integration (CI), .78 for Cross-Source Inquiry (CSI), and .84 for Reflective Calibration (RC), all exceeding the recommended threshold of .70 (Hair et al., 2009). These findings suggest that the AE-AI Scale and its subscales are highly reliable.

4.5. Criterion validity

AE-AI Scale scores demonstrated acceptable criterion validity, as they were positively associated with AI literacy. Specifically, AE-AI Scale scores were strongly correlated with AILS scores ($r = .62$, $p < .001$), indicating that students who reported higher levels of agentic engagement with AI also reported higher AI literacy. Pairwise correlations among the four AE-AI Scale factors were all significant ($p < .01$), ranging from small ($r = .23$) to large ($r = .66$) correlations, as shown in Table 5. Each factor was significantly related to AI literacy ($p < .01$), with correlations ranging from moderate ($r = .37$) to large ($r = .54$). Taken together, these findings support the criterion validity of the AE-AI Scale, demonstrating that agentic engagement with AI is meaningfully associated with established

Table 2
EFA results.

Dimension	Item	Factor				Communality	Uniqueness
		1	2	3	4		
Adaptive Direction (AD)	AD1	**0.73**	.07	.04	.08	.69	.31
	AD2	**1.00**	.02	−.05	−.07	.91	.09
	AD3	**0.78**	.03	.04	.08	.73	.27
	AD4	**0.78**	−.09	.12	.09	.74	.26
Critical Integration (CI)	CI1	.26	.19	**0.43**	.05	.57	.43
	CI2	.22	.04	**0.46**	.05	.44	.56
	CI3	.02	.04	**0.85**	.07	.84	.16
	CI4	.28	.22	**0.55**	−.06	.71	.29
Cross-Source Inquiry (CSI)	CSI 1	.14	.16	−.08	**0.52**	.44	.56
	CSI 2	.07	.13	−.19	**0.73**	.58	.42
	CSI 3	.07	−.04	.27	**0.67**	.69	.31
	CSI 4	.05	.05	.29	**0.58**	.67	.33
Reflective Calibration (RC)	RC1	−.15	**0.54**	.16	.34	.64	.36
	RC2	−.01	**0.84**	.08	−.05	.71	.29
	RC3	.05	**0.82**	−.09	.02	.67	.33
	RC4	.07	**0.67**	.06	.07	.61	.39

Note. Extraction method = maximum likelihood, rotation method = oblique.

Table 3
EFA interfactor correlations.

	AD	CI	CSI	RC
AD	1			
CI	.45	1		
CSI	.58	.44	1	
RC	.40	.59	.45	1

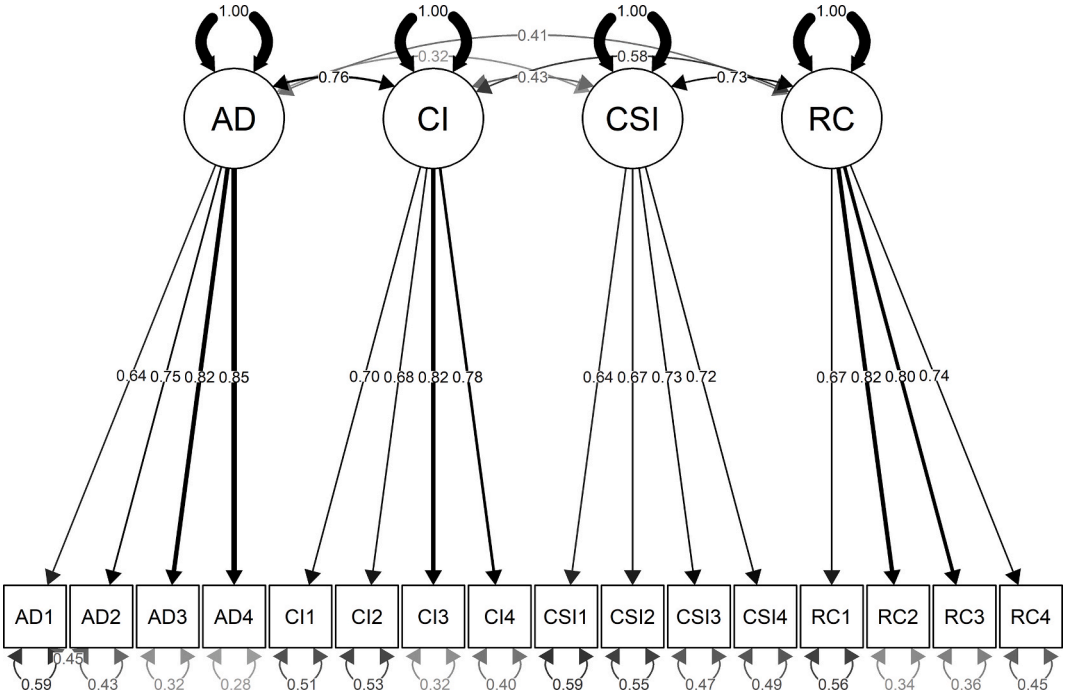


Fig. 3. The four-factor structure model of the AE-AI Scale.
Note. All factor loadings are significant at $p < .001$.

Table 4
Descriptive statistics and reliability analysis (N = 256).

	M	*SD*	Cronbach's α	Composite reliability
AD	4.95	.68	.87	.85
CI	4.92	.68	.83	.84
CSI	4.52	.84	.78	.78
RC	4.66	.76	.84	.84
AE-AI (total)	4.76	.57	.89	N/A

Note. *M* = mean, *SD* = standard deviation, N/A = Not Applicable.

measures of AI literacy and related constructs.

Table 6 presents the HTMT ratios used to assess discriminant validity between the four AE-AI Scale factors and the four AILS dimensions. All cross-construct HTMT values ranged from .25 to .73, remaining below the conservative threshold of .85 (Henseler et al., 2015). The highest HTMT value was observed between AE-AI_CI and AILS_Evaluation (HTMT = .73), which is still within the acceptable range. This pattern indicates satisfactory discriminant validity, suggesting that the AE-AI Scale factors are empirically distinguishable from AI literacy despite their expected conceptual relatedness.

5. Conclusion and discussion

The development and validation of the Agentic Engagement with AI (AE-AI) Scale provide a timely response to the growing attention to student agency in the era of GenAI (Dai et al., 2023; Darvishi et al., 2024). The AE-AI Scale, grounded in the agentic engagement theory, has challenged the traditional unidirectional view of agency and instead framed student agency as a proactive,

Table 5
Correlation analysis between variables (N = 256).

	1	2	3	4	5
1. AE-AI_AD	1				
2. AE-AI_CI	.66**	1			
3. AE-AI_CSI	.23**	.32**	1		
4. AE-AI_RC	.40**	.51**	.59**	1	
5. AILS	.54**	.52**	.37**	.49**	1

Note. **. The correlation is significant at the $p < .01$ level (2-tailed).

reciprocal, and transactional process enacted through ongoing exchanges between human intentionality and machine adaptivity. Through EFA and CFA, the study verified a four-factor structure of the AE-AI Scale: Adaptive Direction (AD), Critical Integration (CI), Cross-Source Inquiry (CSI), and Reflective Calibration (RC). Each factor represents a particular type of strategies that enable students to exercise agency when learning with AI.

Adaptive Direction (AD) captures how students iteratively compose, test, and refine their instructions to shape GenAI outputs in directions useful for their learning. The items describe a recursive cycle: the students give instruction (e.g., "I give clear and detailed instructions to AI to obtain the output I need"), observes the outputs, adjusts, and tries again (e.g., "I explore different ways to instruct AI until I get satisfactory outputs"). Three of the four items refer to this adjustment directly, including tailoring instructions, exploring different ways to instruct, and adding information when outputs fall short. The defining quality of this factor, therefore, goes beyond initial goal-setting, but the capability to adapt human input in response to machine output while maintaining a learning goal (Chen et al., 2024).

This adaptive quality extends conventional self-regulated (Winne & Perry, 2000; Zimmerman, 2002) or learner autonomy theories (Boud, 2012; Little, 2007), which already recognize planning, monitoring, feedback, and strategic adjustment in relatively stable learning conditions. GenAI introduces a more dynamic interactional condition: student agency cannot rest on forethought or planned control, but also require real-time adjustment as students interpret AI responses and decide how to redirect the interaction. The AD factor therefore extends the notion of agency in two ways. First, it frames agency as iterative and recursive rather than one-shot: the student cycles between instruction and output, treating each AI response as feedback for the next move. Second, it situates agency within a human–AI interactional loop: the student remains the intentional actor, but the expression of agency is shaped by what the AI produces.

Critical Integration (CI) captures how students evaluate AI outputs and rework them through their own knowledge and judgement. The items describe a two-step posture: the student first assesses the quality and relevance of AI outputs (e.g., "I assess the quality of AI outputs to ensure they align with my needs"), then selects, enhances, or modifies them to fit the specific demands of learning tasks (e.g., "I modify AI outputs to fit the specific contexts of my tasks"). This factor highlights students' *active reworking* of AI-generated contents through their own intellectual contribution. AI outputs serve as raw material rather than finished products, and the student remains accountable for the final intellectual product.

This factor links agency to *epistemic responsibility*: students take ownership of what counts as valid, relevant, and useable knowledge in their work, rather than deferring to the machine. This resonates with Matusov et al. (2016)'s notion of authorial agency and Wu et al. (2025)'s account of epistemic agency, both of which emphasize the student's role as an active knowledge creator rather than a passive consumer of given content (Lai & Dai, 2025). Such epistemic responsibility also serves as a counterweight to AI limitations: by filtering and modifying outputs, students respond to known weaknesses of GenAI, including inaccuracy, generic phrasing, and context-insensitivity (Yusuf et al., 2024). This may help explain the close association between agentic engagement with AI literacy, since the two constructs reinforce each other: AI literacy supports students' evaluation of outputs, while critical integration gives students opportunities to practice and deepen that literacy in use (Arseven & Bal, 2025; Walter, 2024).

Cross-Source Inquiry (CSI) captures students' deliberate efforts to extend their AI-assisted learning by drawing on sources beyond the AI systems. The items describe two complementary strategies: seeking human input from peers, mentors, and experts (e.g., "I seek guidance from peers or mentors if I find gaps in what AI provides") and consulting external materials such as textbooks, academic sources, or other reference content (e.g., "I compare AI outputs with other materials to increase the depth of my learning"). The core of this factor is that students do not treat AI as a closed loop or final authority; instead, they situate the human-AI loop in the broader learning ecology and actively leverage that ecology to verify, supplement, and deepen their AI-assisted work. While this practice is compatible with AI literacy in recognizing AI's limits, it foregrounds a distinct form of agency: strategically orchestrating multiple sources of expertise to improve learning quality.

Conceptually, CSI aligns with views of agency as distributed across people, tools, and materials**,** rather than located solely within an

Table 6
HTMT ratios for discriminant validity among AE-AI and AILS.

	AILS_Awareness	AILS_Usage	AILS_Evaluation	AILS_Ethics
AE-AI_AD	.54	.61	.64	.31
AE-AI_CI	.49	.58	.73	.28
AE-AI_CSI	.33	.25	.37	.48
AE-AI_RC	.47	.41	.63	.39

individual student (Mairitsch et al., 2023). It also resonates with recent notions of co-agency (Nordström et al., 2023; Roe & Perkins, 2024), distributed agency (Rammert, 2025), and relational agency (Edwards & D'arcy, 2004), which emphasize that learning and decision-making emerge through coordination among multiple actors and mediating resources. From this perspective, student agency is enacted through relational and distributed practices, e.g., negotiating meanings, integrating diverse inputs, rather than through isolated control over a single tool (Mouta et al., 2025).

Reflective Calibration (RC) captures students' ongoing self-examination of their AI-assisted learning and their deliberate adjustment of future AI use based on that reflection. Students reflect on both their own intentions and the tool's contribution (e.g., "I think about whether my AI use serves my learning goals") and translate these reflections into subsequent adjustments in use (e.g., "I evaluate the strengths and weaknesses of AI-aided tasks to adjust my GenAI use"). The defining feature of this factor is not reflection per se, but reflection coupled with calibration: students enhance the loop by using reflective insights to refine their engagement with AI over time.

The RC factor partly aligns with metacognition and self-regulated learning, where monitoring and reflection support strategic adjustment (Fan et al., 2025; Zimmerman, 2002). However, it extends these traditions by directing reflection towards both the students and the tool. In AI-assisted learning, students must evaluate not only their own approaches but also the affordances, limits, and implications of the AI system they are engaging with. This makes reflection prospective and action-oriented: it supports iterative improvement rather than post-hoc evaluation alone. Framed within agentic engagement, RC highlights how students intentionally shape the trajectory of human–AI collaboration through ongoing adjustment, strengthening their adaptability and responsibility in AI-assisted learning (Yang et al., 2024).

6. Limitations and future studies

This study has several limitations, opening avenues for further research. First, this study generated initial items based on student interviews, without involving instructors and experts at this phase. Future research could complement student interviews with expert interviews during early-stage item development. The second limitation concerns potential bias generated by the nature of the scale and the sample. As a self-report Likert instrument, responses may reflect participants' subjective perceptions. In addition, students who chose to participate may hold relatively favorable attitudes toward AI, which could introduce selection bias. This subjectivity also reflects the context-sensitive nature of agency, as social norms and policy environments can shape individuals' understandings of it. As such, cross-contextual validation of the AE-AI Scale is essential to examine its applicability across different educational and cultural settings. Last but not least, while the present study focused on general-purpose AI tools in higher education, future research should examine how emerging forms of AI and human–AI interaction, such as multimodal interactions and embodied agents, may reshape expressions of student agency. As AI technologies rapidly evolve, it is important to assess the AE-AI Scale's continued relevance and adaptability across diverse educational and technological contexts.

CRediT authorship contribution statement

Yun Dai: Writing – review & editing, Writing – original draft, Supervision, Resources, Methodology, Investigation, Funding acquisition, Conceptualization. **Suya Liu:** Software, Formal analysis, Data curation. **Sihan Zhou:** Writing – review & editing, Validation, Resources. **Sichen Lai:** Formal analysis, Data curation. **Ang Liu:** Writing – review & editing, Resources, Data curation. **Cher Ping Lim:** Writing – review & editing, Supervision, Conceptualization.

Ethics statement

This research was granted ethical approval by the Survey and Behavioral Research Ethics Committee at the Chinese University of Hong Kong and by the ethics committees of two participating universities. Informed consent was obtained from all individual participants included in the research.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used ChatGPT and Grammarly in order to improve the language. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Funding information

N.A.

Declaration of competing interest

The authors have no conflicts of interest to disclose.

Appendix A. Item refinement and deletion based on expert review and EFA analysis

Initial item	Primary origin	Action from expert review	Action from EFA
1. I give clear and detailed instructions to AI to obtain the output I need.	Inductive + Deductive (Reeve, 2013)		
2. I tailor and adjust my instructions to ensure that AI generates what I want.	Inductive		
3. I explore different ways to instruct AI until I get satisfactory outputs.	Inductive + Deductive (Mameli & Passini, 2019)		
4. I [~~feed~~ + give] additional information to AI to make its outputs more aligned with my goals.	Inductive	Revision	
5. I provide feedback to AI to improve its performance.	Inductive + Deductive (Reeve, 2013)	Removal (CVI <1)	
6. I guide AI to change or correct its content when its output does not fulfill my needs.	Inductive		Removal (factor loading < |.40|)
7. I adjust the settings of AI to improve the alignment between its inputs and my needs.	Inductive		Removal (communality <.3)
8. I think about whether AI outputs match the purpose of my task before using them.	Inductive		Removal (factor loading < |.40|)
9. I assess the quality of AI outputs to ensure they align with my needs.	Inductive		
10. I weigh the [~~pro and cons~~ strength and weakness] of including AI outputs in my work.	Inductive	Revision	Removal (factor loading < |.40|)
11. I only choose parts of AI outputs that fit my goals and needs.	Inductive		
12. I discard AI outputs when they do not meet my standards or requirements.	Inductive + Deductive (Mameli & Passini, 2019)		Removal (communality <.3)
13. I enhance the quality of AI outputs with my knowledge and judgement.	Inductive		
14. I modify AI outputs to fit the specific contexts of my tasks.	Inductive + Deductive (Kim & Song, 2023)		
15. I consult additional resources when I have doubts about AI outputs.	Inductive	Removal (CVI <1)	
16. I [~~use trusted resources to cross-check AI outputs~~ + cross-check AI outputs with trusted resources].	Inductive	Revision	Removal (cross-loading on two factors >.3 and differences between them <.15)
17. I search for external information to confirm the reliability of AI outputs.	Inductive		Removal (factor loading < |.40|)
18. I discuss my AI-integrated work with others to improve and refine it.	Inductive + Deductive (Kim & Song, 2023)		
19. I seek guidance from peers or mentors if I find gaps in what AI provides.	Deductive (Kim & Song, 2023)		
20. I look for expert opinions [+ or alternative sources] when AI is unable to help me.	Inductive + Deductive (Kim & Song, 2023)	Revision	
21. I compare AI outputs with other materials to increase the depth of my learning.	Inductive + Deductive (Kim & Song, 2023)		
22. I examine my interactions with AI to determine if they are effective.	Inductive + Deductive (Dai & Lai, 2025)		Removal (cross-loading on two factors >.3 and differences between them <.15)
23. I think about whether my AI use serves my learning goals.	Inductive		
24. I consider how I could improve my interactions with AI for better outcomes.	Inductive	Removal (CVI <1)	
25. I evaluate the strengths and weaknesses of AI-aided tasks to adjust my GenAI use.	Inductive		
26. I revisit my AI-aided tasks to identify [~~useful lessons~~ *lessons that can guide future actions*].	Inductive + Deductive (Dai & Lai, 2025)	Revision	
27. I learn from unsatisfying interactions with AI.	Inductive	Removal (CVI <1)	
28. I monitor my task progress to see if AI supports my learning.	Deductive (Dai & Lai, 2025)		

Note. Revisions are shown in square brackets [ ] in the column "Initial item". Explanations for deletions are shown in parentheses () in the columns "Action from expert review" and "Action from EFA".

Appendix B. Sample item using the 6-point Likert scale

	Strongly disagree	Disagree	Slightly disagree	Slightly agree	Agree	Strongly Agree
I give clear and detailed instructions to AI to obtain the output I need.	1	2	3	4	5	6

Appendix C. Factors and items for the Agentic Engagement with AI (AE-AI) Scale

Factors & Items	*Mean*	*SD*
Factor 1: Adaptive Direction (AD)		
1. I give clear and detailed instructions to AI to obtain the output I need.	4.86	.83
2. I tailor and adjust my instructions to ensure that AI generates what I want.	4.91	.76
3. I explore different ways to instruct AI until I get satisfactory outputs.	5.00	.81
4. I give additional information to AI to make its outputs more aligned with my goals.	5.04	.80
Factor 2: Critical Integration (CI)		
5. I assess the quality of AI outputs to ensure they align with my needs.	4.89	.84
6. I only choose parts of AI outputs that fit my goals and needs.	4.87	.88
7. I enhance the quality of AI outputs with my knowledge and judgement.	4.98	.81
8. I modify AI outputs to fit the specific contexts of my tasks.	4.94	.79
Factor 3: Cross-Source Inquiry (CSI)		
9. I discuss my AI-integrated work with others to improve and refine it.	4.31	1.15
10. I seek guidance from peers or mentors if I find gaps in what AI provides.	4.27	1.27
11. I look for expert opinions or alternative sources when AI is unable to help me.	4.72	.96
12. I compare AI outputs with other materials to increase the depth of my learning.	4.77	.89
Factor 4: Reflective Calibration (RC)		
13. I think about whether my AI use serves my learning goals.	4.77	.84
14. I evaluate the strengths and weaknesses of AI-aided tasks to adjust my GenAI use.	4.71	.91
15. I revisit my AI-aided tasks to identify lessons that can guide future actions.	4.55	1.01
16. I monitor my task progress to see if AI supports my learning.	4.60	.97

Data availability

Data will be made available on request.

References

Anton, J., Cosentino, G., Sharma, K., Gelsomini, M., Mok, M., Giannakos, M., & Abrahamson, D. (2025). The human condition: Modal and interactive advantages of teacher over AI feedback on children's mathematical performance. In *Proceedings of the 24th interaction design and children* (pp. 183–203).

Antonietti, C., Schmitz, M.-L., Consoli, T., Cattaneo, A., Gonon, P., & Petko, D. (2023). Development and validation of the ICAP technology scale to measure how teachers integrate technology into learning activities. *Computers & Education, 192*. https://doi.org/10.1016/j.compedu.2022.104648

Arseven, T., & Bal, M. (2025). Critical literacy in artificial intelligence assisted writing instruction: A systematic review. *Thinking Skills and Creativity, 57*, Article 101850.

Avsec, S., & Rupnik, D. (2025). From transformative agency to AI literacy: Profiling Slovenian technical high school students through the five big ideas lens. *Systems, 13*(7), 562.

Bandura, A. (2014). Exercise of personal agency through the self-efficacy mechanism. In *Self-efficacy* (pp. 3–38). Taylor & Francis.

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health, 6*, 149.

Bontcheva, K., Papadopoulous, S., Tsalakanidou, F., Gallotti, R., Krack, N., Teyssou, D., France-Presse, A., Cuccovillo, L., & Verdoliva, L. (2024). Generative AI and disinformation: Recent advances, challenges, and opportunities [White paper]. *European Digital Media Observatory*. https://edmo.eu/wp-content/uploads/2023/12/Generative-AI-and-Disinformation_-White-Paper-v8.pdf.

Boud, D. (2012). *Developing student autonomy in learning*. Routledge.

Brod, G., Kucirkova, N., Shepherd, J., Jolles, D., & Molenaar, I. (2023). Agency in educational technology: Interdisciplinary perspectives and implications for learning design. *Educational Psychology Review, 35*(1), 25.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. Guilford publications.

Chen, A., Lyu, A., & Lu, Y. (2024). Member's performance in human–AI hybrid teams: A perspective of adaptability theory. *Information Technology & People, 39*(1), 157–177.

Chomeya, R. (2010). Quality of psychology test between Likert scale 5 and 6 points. *Journal of Social Sciences, 6*(3), 399–403.

Code, J. (2020). Agency for learning: Intention, motivation, self-efficacy and self-regulation. *Frontiers in Education, 5*, 19. https://doi.org/10.3389/feduc.2020.00019

Code, J. (2025). The entangled learner: Critical agency for the postdigital era. *Postdigital Science and Education, 7*, 336–358. https://doi.org/10.1007/s42438-025-00544-1

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge.

Cohen, R., Moed, A., Shoshani, A., Roth, G., & Kanat-Maymon, Y. (2020). Teachers' conditional regard and students' need satisfaction and agentic engagement: A multilevel motivation mediation model. *Journal of Youth and Adolescence, 49*, 790–803.

Cosentino, G., Anton, J., Sharma, K., Gelsomini, M., Giannakos, M., & Abrahamson, D. (2025). Bridging AI and human feedback: Hybrid intelligence in embodied math education. In *Joint proceedings of LAK 2025 workshops, co-located with the 15th International Conference on Learning Analytics and Knowledge (LAK 2025)*, Dublin, Ireland, March 3–7, 2025.

Costello, A. B., & Osborne. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research and Evaluation, 10*(1). https://doi.org/10.7275/jyj1-4868

Cox, B. E., McIntosh, K., Reason, R. D., & Terenzini, P. T. (2014). Working with missing data in higher education research: A primer and real-world example. *The Review of Higher Education, 37*(3), 377–402. https://doi.org/10.1353/rhe.2014.0026

Dai, Y., Lai, S., Lim, C. P., & Liu, A. (2023). ChatGPT and its impact on research supervision: Insights from Australian postgraduate research students. *Australasian Journal of Educational Technology, 39*(4), 74–88.

Dai, Y., Liu, A., & Lim, C. P. (2023). Reconceptualizing ChatGPT and generative AI as a student-driven innovation in higher education. *Procedia CIRP, 119*, 7. https://doi.org/10.1016/j.procir.2023.05.002

Dai, Y., & Lai, S. (2025). A theoretical framework of student agency in AI-assisted learning: A grounded theory approach. In K. Machtmes, M. L. Wu, & L. Khan (Eds.), *Oxford Intersections: Social Media in Society and Culture*. Oxford University Press. https://doi.org/10.1093/9780198945253.003.0009.
Darvishi, A., Khosravi, H., Sadiq, S., Gasevic, D., & Siemens, G. (2024). Impact of AI assistance on student agency. *Computers & Education, 210*. https://doi.org/10.1016/j.compedu.2023.104967. Article 104967.
Dörnyei, Z., & Dewaele, J.-M. (2022). *Questionnaires in second Language research: Construction, administration, and processing*. Taylor & Francis.
Edwards, A., & D'arcy, C. (2004). Relational agency and disposition in sociocultural accounts of learning to teach. *Educational Review, 56*(2), 147–155.
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.
Finn, J. D., & Zimmer, K. S. (2012). Student engagement: What is it? Why does it matter?. In *Handbook of research on student engagement* (pp. 97–131). Springer.
Floridi, L. (2025). AI as agency without intelligence: On artificial intelligence as a new form of artificial agency and the multiple realisability of agency thesis: L. floridi. *Philosophy & Technology, 38*(1), 30.
Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59–109.
George, A. S. (2023). The potential of generative AI to reform graduate education. *Partners Universal International Research Journal, 2*(4), 36–50.
Guo, W., Lianga, Z., Wang, C., Li, X., Hu, H., Chen, S., Yua, Q., & Zhaoa, Q. (2025). Student-AI collaborative creative Problem-Solving: The role of human agency. *Computers & Education*, Article 105433.
Gupta, N., Ali, K., Jiang, D., Fink, T., & Du, X. (2024). Beyond autonomy: Unpacking self-regulated and self-directed learning through the lens of learner agency-a scoping review. *BMC Medical Education, 24*(1), 1519.
Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). *Multivariate data analysis* (7th ed.). Prentice-Hall.
Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*(1), 115–135.
Holgado-Tello, F. P., Chacón-Moscoso, S., Barbero-García, I., & Vila-Abad, E. (2008). Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. *Quality and Quantity, 44*(1), 153–166. https://doi.org/10.1007/s11135-008-9190-y
Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education, 57*(4), 542–570.
Holstein, K., McLaren, B. M., & Aleven, V. (2019). Designing for complementarity: Teacher and student needs for orchestration support in AI-enhanced classrooms. In *International conference on artificial intelligence in education*.
Horn, J. L. (1965). A rationale and Test for the number of factors in factor analysis. *Psychometrika, 30*, 179–185. https://doi.org/10.1007/BF02289447
Hou, C., Zhu, G., Sudarshan, V., Lim, F. S., & Ong, Y. S. (2025). Measuring undergraduate students' reliance on Generative AI during problem-solving: Scale development and validation. *Computers & Education, 234*. https://doi.org/10.1016/j.compedu.2025.105329
Iacobucci, D., Ruvio, A., Román, S., Moon, S., & Herr, P. M. (2022). How many factors in factor analysis? New insights about parallel analysis with confidence intervals. *Journal of Business Research, 139*, 1026–1043. https://doi.org/10.1016/j.jbusres.2021.09.015
Jääskelä, P., Poikkeus, A.-M., Vasalampi, K., Valleala, U. M., & Rasku-Puttonen, H. (2017). Assessing agency of university students: Validation of the AUS Scale. *Studies in Higher Education, 42*(11), 2061–2079.
James, T. L., Zhang, J., Li, H., Ziegelmayer, J. L., & Villacis-Calderon, E. D. (2022). The moderating effect of technology overload on the ability of online learning to meet students' basic psychological needs. *Information Technology & People, 35*(4), 1364–1382.
Jiang, J., Vetter, M. A., & Lucia, B. (2024). Toward a 'more-than-digital'AI literacy: Reimagining agency and authorship in the postdigital era with ChatGPT. *Postdigital Science and Education, 6*(3), 922–939.
Khine, M. S. (2024). Using AI for adaptive learning and adaptive assessment. In *Artificial intelligence in education: A machine-generated literature overview* (pp. 341–466). Springer.
Kim, H. Y. (2013). Statistical notes for clinical researchers: Assessing normal distribution (2) using skewness and kurtosis. *Restor Dent Endod, 38*(1), 52–54. https://doi.org/10.5395/rde.2013.38.1.52
Kim, R., & Song, H.-D. (2023). Developing an agentic engagement scale in a self-paced MOOC. *Distance Education, 44*(1), 120–136.
Klemenčič, M. (2017). From student engagement to student agency: Conceptual considerations of European policies on student-centered learning in higher education. *Higher Education Policy, 30*, 69–85.
Klemenčič, M. (2020). Students as actors and agents in student-centered higher education. In *The Routledge international handbook of student-centered learning and teaching in higher education* (pp. 92–108). Routledge.
Kline, R. B. (2023). *Principles and practice of structural equation modeling*. Guilford publications.
Koenigs, S. S., Fiedler, M. L., & Decharms, R. (1977). Teacher beliefs, classroom interaction and personal causation 1. *Journal of Applied Social Psychology, 7*(2), 95–114.
Kostopoulos, G., Gkamas, V., Rigou, M., & Kotsiantis, S. (2025). Agentic AI in education: State of the art and future directions, *IEEE Access*,13, 177467 - 177491.
Lai, S., & Dai, Y. (2025). Agentic scientists or exhausted workers? Supervision, identity, and epistemic agency of medical doctoral students in Hong Kong. *International Journal of Educational Management*, 1–17.
Lai, S., Liu, S., Dai, Y., Liu, A., & Lim, C. P. (2025). The impacts and tensions of generative AI on doctoral students' supervisory and peer dynamics: An activity theory analysis. *Australasian Journal of Educational Technology, 41*(5), 1–17.
Leighton, J. P. (2017). *Using think-aloud interviews and cognitive labs in educational research*. Oxford University Press.
Little, D. (2007). Language learner autonomy: Some fundamental considerations revisited. *International Journal of Innovation in Language Learning and Teaching, 1*(1), 14–29.
Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI conference on human factors in computing systems*.
Mahdavi Goloujeh, A., Sullivan, A., & Magerko, B. (2024). Is it AI or is it me? Understanding users' prompt journey with text-to-image generative AI tools. In *Proceedings of the 2024 CHI conference on human factors in computing systems*.
Mairitsch, A., Sulis, G., Mercer, S., & Bauer, D. (2023). Putting the social into learner agency: Understanding social relationships and affordances. *International Journal of Educational Research, 120*, Article 102214.
Mameli, C., & Passini, S. (2019). Development and validation of an enlarged version of the student agentic engagement scale. *Journal of Psychoeducational Assessment, 37*(4), 450–463.
Marín, V. I., de Benito, B., & Darder, A. (2024). Technology-enhanced learning for student agency in higher education: A systematic literature review. *Interaction Design and Architecture (s) Journal-IxD&A, 2020*(45), 15–49.
Matusov, E., Smith, M., Soslau, E., Marjanovic-Shane, A., & Von Duyke, K. (2016). Dialogic education for and from authorial agency. *Dialogic Pedagogy, 4*.
McGrath, C., Farazouli, A., & Cerratto-Pargman, T. (2024). Generative AI chatbots in higher education: A review of an emerging research area. *Higher Education*, 1–17.
Molenaar, I. (2022). The concept of hybrid human-AI regulation: Exemplifying how to support young learners' self-regulated learning. *Computers and Education: Artificial Intelligence, 3*, Article 100070.
Mouta, A., Pinto-Llorente, A. M., & Torrecilla-Sánchez, E. M. (2025). "Where is agency moving to?": Exploring the interplay between AI technologies in education and human agency. *Digital Society, 4*(2), 49.
Ng, D. T. K., Leung, J. K. L., Chu, K. W. S., & Qiao, M. S. (2021). AI literacy: Definition, teaching, evaluation and ethical issues. *Proceedings of the Association for Information Science and Technology, 58*(1), 504–509.
Nopas, D.-s. (2025). Algorithmic learning or learner autonomy? Rethinking AI's role in digital education. *Qualitative Research Journal, Online Version*.
Nordström, P., Lundman, R., & Hautala, J. (2023). Evolving coagency between artists and AI in the spatial cocreative process of artmaking. *Annals of the American Association of Geographers, 113*(9), 2203–2218.
Nwana, H. S. (1990). Intelligent tutoring systems: An overview. *Artificial Intelligence Review, 4*(4), 251–277.

Padiyath, A., Guzdial, M., & Ericson, B. (2025). Development of the critical reflection and agency in Computing index. In *Proceedings of the 2025 CHI conference on human factors in computing systems*.
Pineda-Báez, C., Manzuoli, C. H., & Sánchez, A. V. (2019). Supporting student cognitive and agentic engagement: Students' voices. *International Journal of Educational Research, 96*, 81–90.
Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health, 29*(5), 489–497.
Rammert, W. (2025). Technology and the Theory of Distributed Agency: A New Approach After Actor-Network-Theory. In *Distributed Agency & Digital Technology: A Social Pragmatist View on Human-Technology Interaction* (pp. 129-142). Wiesbaden: Springer Fachmedien Wiesbaden.
Reeve, J. (2012). A self-determination theory perspective on student engagement. In *Handbook of research on student engagement* (pp. 149–172). Springer.
Reeve, J. (2013). How students create motivationally supportive learning environments for themselves: The concept of agentic engagement. *Journal of Educational Psychology, 105*(3), 579.
Reeve, J., Cheon, S. H., & Yu, T. H. (2020). An autonomy-supportive intervention to develop students' resilience by boosting agentic engagement. *International Journal of Behavioral Development, 44*(4), 325–338.
Reeve, J., & Shin, S. H. (2020). How teachers can support students' agentic engagement. *Theory into practice, 59*(2), 150–161.
Reeve, J., & Tseng, C.-M. (2011). Agency as a fourth aspect of students' engagement during learning activities. *Contemporary Educational Psychology, 36*(4), 257–267.
Rodríguez-Castro, Y., Alonso-Ruido, P., González-Fernández, A., Lameiras-Fernández, M., & Carrera-Fernández, M. V. (2017). Spanish adolescents' attitudes towards sexting: Validation of a scale. *Computers in Human Behavior, 73*, 375–384. https://doi.org/10.1016/j.chb.2017.03.049
Roe, J., & Perkins, M. (2024). Generative AI and agency in Education: A critical scoping review and thematic analysis. Preprint version, https://arxiv.org/abs/2411.00631.
Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching, 6*(1).
Saini, C., & Abraham, J. (2021). Development and psychometric evidence of an engagement instrument in the pre-service teacher education context. *Studies In Educational Evaluation, 71*. https://doi.org/10.1016/j.stueduc.2021.101095
Satyanarayan, A., & Jones, G. M. (2024). Intelligence as agency: Evaluating the capacity of generative AI to empower or constrain human action. *An MIT exploration of generative AI*. https://doi.org/10.21428/e4baedd9.2d7598a2
Schmitt, T. A., & Sass, D. A. (2011). Rotation criteria and Hypothesis testing for exploratory factor analysis: Implications for factor pattern loadings and interfactor correlations. *Educational and Psychological Measurement, 71*(1), 95–113. https://doi.org/10.1177/0013164410387348
Scott, W. R. (2008). *Institutions and organizations: Ideas and interests*. Sage.
Shogren, K. A., Wehmeyer, M. L., Palmer, S. B., Forber-Pratt, A. J., Little, T. J., & Lopez, S. (2015). Causal agency theory: Reconceptualizing a functional model of self-determination. *Education and training in autism and developmental disabilities, 50*(3), 251–263.
Singh, K., Junnarkar, M., & Kaur, J. (2016). *Measures of positive psychology: Development and validation*. Springer. https://doi.org/10.1007/978-81-322-3631-3
Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Pearson.
Trowler, V. (2010). Student engagement literature review. *The higher education academy, 11*(1), 1–15.
Uanachain, D. M. N., & Aouad, L. I. (2025). Generative AI in education: Rethinking learning, assessment & student agency for the AI era. *Thresholds in Education, 48*(1), 111–132.
Vincoli, M., Scholz, N., & Nagashima, T. (2025). Multidimensional student agency in learning with AI: A conceptual framework and design implications. In *International conference on artificial intelligence in education*.
Walter, Y. (2024). Embracing the future of Artificial Intelligence in the classroom: The relevance of AI literacy, prompt engineering, and critical thinking in modern education. *International Journal of Educational Technology in Higher Education, 21*(1), 15.
Wang, B., Rau, P.-L. P., & Yuan, T. (2022). Measuring user competence in using artificial intelligence: Validity and reliability of artificial intelligence literacy scale. *Behaviour & Information Technology, 42*(9), 1324–1337. https://doi.org/10.1080/0144929x.2022.2072768
Watkins, M. W. (2018). Exploratory factor analysis: A guide to best practice. *Journal of Black Psychology, 44*(3), 219–246.
Winne, P. H., & Perry, N. E. (2000). Measuring self-regulated learning. In *Handbook of self-regulation* (pp. 531–566). Elsevier.
Wise, A. F., Vytasek, J. M., Hausknecht, S., & Zhao, Y. (2016). Developing learning analytics design knowledge in the" middle space": The student tuning model and align design framework for learning analytics use. *Online Learning, 20*(2), 155–182.
Wu, J.-Y., Lee, Y.-H., Chai, C. S., & Tsai, C.-C. (2025). Strengthening human epistemic agency in the symbiotic learning partnership with generative artificial intelligence. *Educational Researcher*, Article 0013189X251333628.
Wu, M.-J., Zhao, K., & Fils-Aime, F. (2022). Response rates of online surveys in published research: A meta-analysis. *Computers in Human Behavior Reports, 7*, Article 100206.
Xia, L., Shen, K., Sun, H., An, X., & Dong, Y. (2025). Developing and validating the student learning agency scale in generative artificial intelligence (AI)-supported contexts. *Education and Information Technologies*, 1–23.
Yang, Y., Luo, J., Yang, M., Yang, R., & Chen, J. (2024). From surface to deep learning approaches with Generative AI in higher education: An analytical framework of student agency. *Studies in Higher Education*, 1–14.
Yusuf, A., Bello, S., Pervin, N., & Tukur, A. K. (2024). Implementing a proposed framework for enhancing critical thinking skills in synthesizing AI-generated texts. *Thinking Skills and Creativity, 53*, Article 101619.
Zarrinabadi, N., Lou, N. M., & Shirzad, M. (2021). Autonomy support predicts language mindsets: Implications for developing communicative competence and willingness to communicate in EFL classrooms. *Learning and Individual Differences, 86*, Article 101981.
Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., Liu, J.-B., Yuan, J., & Li, Y. (2021). A Review of Artificial Intelligence (AI) in Education from 2010 to 2020. *Complexity, 2021*(1), Article 8812542.
Zhou, S., Xu, J., & Thomas, N. (2024). L2 listening in a digital era: Developing and validating the mobile-assisted self-regulated listening strategy questionnaire (MSRLS-Q). *System, 123*. https://doi.org/10.1016/j.system.2024.103310
Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into practice, 41*(2), 64–70.