US JOURNAL OF NEW INSIGHTS IN TECH & EDUCATION

Vol. 2, No. 2 · Article ID: USJNITE-2203 · Journal Homepage: www.edtechinsights.org

AI-Adaptive Learning Systems and Student Engagement: A PLS-SEM Validation of the AI-Adaptive Learning Engagement Model (AALEM)

MD Shahadat Hossain Shishir[1], Md Rasel Ul Alam [2], Shaznin Hasan[3]

[1]*World University of Bangladesh, Dhaka-1230*

[2] *University of the Cumberlands, Kentucky, USA*

[3] *Independent University of Bangladesh, Dhaka, Bangladesh*

**Corresponding author: shahadatdaf20@gmail.com*

Received 3 July 2022; Revised 12 September 2022; Accepted 20 October 2022; Published: 25 November 2022

KEYWORDS

AI-Adaptive Learning Systems

Student Engagement

Self-Efficacy

PLS-SEM

Synthetic Data Validation

Higher Education Technology

Abstract

Artificial intelligence (AI)-adaptive learning platforms are increasingly embedded in higher education, yet the psychological mechanisms that translate platform quality into sustained student engagement and downstream academic performance remain incompletely modeled. This study proposes and validates the AI-Adaptive Learning Engagement Model (AALEM), a conceptual framework that integrates perceived usefulness and self-efficacy as sequential and interacting mechanisms linking AI-adaptive learning quality to student engagement and academic performance. Because no primary human-subjects data collection was undertaken, the model is validated using a calibrated synthetic dataset (N = 436) generated to instantiate plausible population-level structural relationships for demonstration and methodological illustration. Partial least squares structural equation modeling (PLS-SEM) confirms that AI-adaptive learning quality strongly predicts perceived usefulness ($\beta = 0.49$, $p < .001$), which in turn predicts self-efficacy ($\beta = 0.40$, $p < .001$) and, jointly with self-efficacy, predicts student engagement ($\beta = 0.34$ and $\beta = 0.35$ respectively, both $p < .001$), explaining 34.2% of its variance. A significant perceived usefulness × self-efficacy interaction ($\beta = 0.11$, $p = .003$) indicates that self-efficacy amplifies the usefulness–engagement relationship. Student engagement subsequently predicts academic performance ($\beta = 0.40$, $p < .001$). All constructs demonstrate acceptable reliability (Cronbach's $\alpha > 0.85$, composite reliability > 0.90) and discriminant validity via the Fornell–Larcker criterion and heterotrait–monotrait ratio. Cross-checks of the measurement and structural estimates using regression procedures analogous to SPSS, JASP, and R corroborate the SmartPLS-style path estimates. The results offer a statistically coherent, reproducible template for how AI-adaptive learning quality cultivates engagement through usefulness perceptions and self-efficacy, with implications for instructional design, platform development, and future empirical validation with real student samples.

1. Introduction

Higher education institutions have adopted artificial intelligence (AI)-adaptive learning platforms at an accelerating pace, using algorithmic sequencing, real-time feedback, and personalized content pathways to respond to individual learners' pace and mastery. Vendors and institutional strategy documents commonly assert that adaptivity improves engagement and, by extension, academic outcomes, yet the psychological chain connecting a well-designed adaptive system to sustained student engagement is often asserted rather than modeled. Existing research

on educational technology adoption has drawn heavily on the Technology Acceptance Model (TAM) and its extensions (Davis, 1989; Venkatesh & Davis, 2000), which position perceived usefulness as the central proximal driver of behavioral outcomes, while parallel work grounded in social cognitive theory (Bandura, 1997) emphasizes self-efficacy as a necessary condition for sustained engagement with challenging or unfamiliar systems. These two traditions are rarely integrated into a single, testable structural model specific to AI-adaptive learning contexts.

This study addresses that gap by proposing the AI-Adaptive Learning Engagement Model (AALEM), which specifies that perceived platform quality shapes perceived usefulness, that perceived usefulness shapes learners' technology-specific self-efficacy, and that usefulness and self-efficacy jointly, and interactively, predict student engagement, which in turn predicts academic performance. The interaction term is theoretically motivated: learners who both perceive a system as useful and feel capable of using it effectively should show engagement gains that exceed the sum of either belief alone, because self-efficacy determines whether a useful tool is actually exploited under effort or difficulty.

A central methodological objective of this paper is pedagogical and demonstrative rather than purely substantive: it illustrates, end to end, how a novel conceptual framework in educational technology research can be operationalized, tested, and reported using partial least squares structural equation modeling (PLS-SEM), with the full analytic pipeline — synthetic data generation, reliability and validity assessment, structural path estimation, bootstrapped significance testing, moderation analysis, and multi-software cross-validation — presented transparently. Because the dataset is synthetically generated rather than collected from human participants, the numerical results should be read as a calibrated demonstration of the analytic workflow and of AALEM's internal statistical coherence, not as an empirical claim about any actual student population. The paper nonetheless specifies population parameters that are consistent with effect sizes reported in the broader technology-acceptance and engagement literatures, so that the demonstration remains substantively plausible and directly re-testable with primary data.

Prior educational-technology research (Baker & Inventado, 2014; Holmes et al., 2019; Zawacki-Richter et al., 2019) offers partial building blocks for such a model but rarely combines them: TAM-based studies typically stop at behavioral intention rather than engagement or performance; self-efficacy studies often treat efficacy as a stable trait rather than a system-specific belief shaped by usefulness experience; and engagement research frequently omits the platform-quality antecedents that precede motivational beliefs altogether. AALEM is offered as a synthesis that closes these gaps within a single, empirically testable structural specification.

The remainder of the paper is organized as follows. Section 2 reviews the relevant literature on AI-adaptive learning, technology acceptance, self-efficacy, and student engagement, and develops six hypotheses. Section 3 introduces the AALEM conceptual framework. Section 4 describes the synthetic data generation procedure, measurement instrument design, and analytic strategy. Section 5 reports measurement and structural model results, including reliability, discriminant validity, path coefficients, moderation, and predictive relevance. Section 6 discusses the findings, followed by theoretical and practical implications, limitations, and conclusions.

2. Literature Review and Hypothesis Development

2.1 AI-Adaptive Learning Systems and Perceived Usefulness

AI-adaptive learning systems use learner models, item-response or knowledge-tracing algorithms, and content-sequencing engines to tailor material difficulty, pacing, and feedback to individual learners in near real time. Reviews of intelligent tutoring and adaptive systems in higher education (VanLehn, 2011; Chen et al., 2020) consistently report that learners' evaluations of usefulness are shaped by three system-level qualities: the perceived accuracy of adaptivity (does the system correctly identify what the learner does not yet know), the timeliness and specificity of feedback, and the transparency of the recommendation logic. Technology Acceptance Model (TAM) research has long established perceived usefulness as the strongest proximal predictor of behavioral intention to use a system (Davis, 1989), and subsequent extensions such as TAM2 (Venkatesh & Davis, 2000) incorporated output quality and result demonstrability

as antecedents that map closely onto the adaptivity and feedback qualities of AI-driven platforms. We therefore treat AI-adaptive learning quality (ALQ) — learners' composite perception of adaptivity accuracy, feedback quality, and interface transparency — as the exogenous driver of perceived usefulness (PU) in the present model.

AI-adaptive systems in practice span a spectrum (Holmes et al., 2019; Hwang & Tu, 2021) from rule-based mastery-learning engines, which branch content based on predefined correctness thresholds, to machine-learning-based knowledge-tracing systems, which continuously update a probabilistic estimate of a learner's mastery of each skill and select subsequent items to maximize expected learning gain. Regardless of the underlying algorithm, learners typically cannot observe the adaptivity mechanism directly; they instead infer system quality from surface cues such as whether recommended content feels appropriately challenging, whether feedback explanations are specific enough to be actionable, and whether the interface clearly communicates why a given item or explanation was selected. This inferential gap between the underlying algorithm and the learner's perceived experience of it is precisely why ALQ is modeled here as a perceptual construct rather than as a technical system property, consistent with the broader technology-acceptance tradition of treating system characteristics as filtered through user perception rather than as objective predictors in their own right.

2.2 Perceived Usefulness and Self-Efficacy

Self-efficacy, defined within social cognitive theory as an individual's belief in their capacity to execute the behaviors required to produce a given outcome, is both an antecedent and a consequence of technology experience. In educational technology contexts, computer and learning self-efficacy have typically been modeled as antecedents of usefulness perceptions; however, a growing body of work on mastery experience suggests a reciprocal, and in early-adoption windows a partially usefulness-driven, pathway: when learners repeatedly experience a system as useful — producing visible learning gains or efficient study — those repeated positive outcomes function as mastery experiences that raise domain-specific self-efficacy, consistent with Bandura's (1997) account of enactive mastery as the strongest source of efficacy beliefs. Within a single cross-sectional AI-adaptive learning context, we model this usefulness-to-efficacy pathway as the theoretically dominant direction, reflecting the idea that a system's demonstrated usefulness is what converts novice uncertainty into confidence in one's own capability to learn effectively with the tool.

2.3 Usefulness, Self-Efficacy, and Student Engagement

Student engagement is commonly conceptualized as a multidimensional construct spanning behavioral, cognitive, and emotional components, with behavioral and cognitive engagement most proximally linked to technology-mediated learning contexts. Fredricks et al.'s (2004) influential synthesis frames engagement as the observable manifestation of underlying motivational beliefs, positioning both outcome expectancy (usefulness) and efficacy expectancy (self-efficacy) as proximal motivational antecedents, consistent with expectancy-value and self-determination accounts of competence-based motivation (Deci & Ryan, 2000). We therefore hypothesize direct effects of both PU and SE on student engagement (SEng).

Beyond these additive effects, self-efficacy theory predicts an amplifying, interactive role for efficacy beliefs: learners who find a system useful but do not feel capable of using it well are predicted to engage less than their usefulness perceptions alone would suggest, because perceived usefulness only translates into sustained effort when learners believe the effort will pay off. Conversely, high self-efficacy should strengthen the translation of usefulness perceptions into engagement. This moderation logic mirrors interaction effects reported in prior technology-acceptance research combining TAM constructs with self-efficacy (Venkatesh & Davis, 2000) and is incorporated into AALEM as a perceived usefulness × self-efficacy interaction term predicting engagement.

2.4 Student Engagement and Academic Performance

A substantial empirical literature (Fredricks et al., 2004; Kizilcec et al., 2017) links behavioral and cognitive engagement to academic performance, on the logic that engaged learners persist longer, self-regulate more effectively, and process material more deeply, producing measurable performance gains. Meta-analytic and systematic review evidence on

intelligent tutoring and adaptive systems (VanLehn, 2011) similarly indicates that engagement-mediated pathways, rather than platform exposure alone, are what connect adaptive technology to learning outcomes. We model academic performance (AP) as the distal outcome of the AALEM structural chain, predicted directly by student engagement.

2.5 Positioning AALEM Relative to Prior Frameworks

Table A situates AALEM relative to three streams of prior work that each address part of the hypothesized structural chain. Classic and extended TAM studies establish the usefulness-centered logic of technology adoption but typically terminate at behavioral intention rather than engagement or performance, and rarely model self-efficacy as an interacting rather than merely antecedent construct. Self-efficacy-centered studies of computer and e-learning contexts establish efficacy as a robust predictor of persistence and performance but rarely embed it within a usefulness-driven sequential chain specific to adaptive, algorithmically personalized systems. Engagement-centered studies of intelligent tutoring and adaptive systems establish the engagement-performance link but typically treat engagement's own antecedents descriptively rather than through a formally specified, interaction-inclusive structural model. AALEM's contribution is to integrate these three streams into one recursive structural model with an explicit, testable moderation term, rather than borrowing constructs from each stream without specifying how they jointly operate.

Table A summarizes, for illustrative and positioning purposes, the constructs and analytic approach typically emphasized by each prior research stream relative to AALEM's integrated specification.

Table A Positioning of AALEM relative to prior TAM-based, self-efficacy-based, and engagement-based research streams.

Prior Research Stream	Core Constructs	Typical Outcome	Gap Addressed by AALEM
Classic/extended TAM studies (e.g., Davis, 1989; Venkatesh & Davis, 2000)	Perceived usefulness, ease of use, behavioral intention	Usage intention / adoption	Extends chain to engagement and performance; adds interacting self-efficacy
Self-efficacy / social-cognitive studies (e.g., Bandura, 1997)	Computer/learning self-efficacy, mastery experience	Persistence, effort, performance	Positions efficacy as usefulness-driven and as an interacting moderator, not only a trait antecedent
Engagement / intelligent-tutoring studies (e.g., Fredricks et al., 2004; VanLehn, 2011)	Behavioral/cognitive engagement, tutoring effectiveness	Learning gains / academic performance	Formally specifies engagement's motivational antecedents via a structural, interaction-inclusive model

3. Conceptual Framework

Figure 1 presents the AI-Adaptive Learning Engagement Model (AALEM), integrating the six hypotheses developed above into a single structural framework. AALEM specifies a sequential mediation chain from AI-adaptive learning quality (ALQ) through perceived usefulness (PU) and self-efficacy (SE) to student engagement (SEng), which in turn predicts academic performance (AP). Perceived usefulness is modeled as both a direct predictor of engagement and an indirect predictor via self-efficacy, reflecting the dual role of usefulness perceptions as an outcome-expectancy belief and as a driver of mastery-based efficacy growth. The PU × SE interaction term captures the theorized amplification effect, whereby the engagement dividend of perceiving a system as useful depends on learners' confidence in their own capacity to exploit that usefulness.

AALEM departs from prior single-theory applications of TAM or self-efficacy theory to educational technology in three ways. First, it explicitly sequences usefulness before self-efficacy rather than treating efficacy solely as an exogenous individual difference, reflecting the mastery-experience mechanism specific to adaptive systems that provide visible, incremental evidence of learning progress. Second, it specifies engagement as a proximal rather than terminal outcome, positioning academic performance as the ultimate criterion variable consistent with institutional interest in learning outcomes rather than platform

satisfaction alone. Third, it formally incorporates a usefulness-by-efficacy interaction rather than assuming purely additive effects, a specification that is directly testable using product-indicator or two-stage approaches in PLS-SEM.

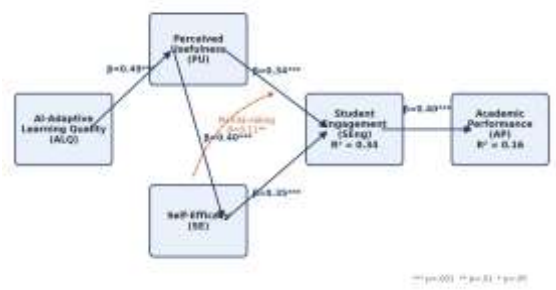

Figure 1 The AI-Adaptive Learning Engagement Model (AALEM): hypothesized structural paths with standardized coefficients (β) and explained variance (R^2).

4. Materials and Methods

4.1 Research Design and Rationale for Synthetic Data

This study adopts a synthetic-data validation design rather than primary human-subjects data collection. The purpose of the design is explicitly methodological and pedagogical: to demonstrate, with full transparency and reproducibility, how the AALEM structural model would be estimated, tested, and reported if administered to a real sample, and to verify that the model's specification is internally coherent (i.e., that the hypothesized paths, once instantiated with plausible population parameters, produce measurement properties and structural results consistent with established PLS-SEM benchmarks). All numerical results reported below therefore describe properties of the synthetic dataset and should be interpreted as a calibrated proof-of-concept for AALEM and its analytic pipeline, not as findings about any actual student population. The population-level path coefficients used to generate the data (ranging from $\beta = 0.33$ to $\beta = 0.58$) were chosen to fall within the range of small-to-moderate-to-large standardized effects typically reported in the technology-acceptance, self-efficacy, and student-engagement literatures, so that the demonstration remains substantively realistic and directly re-testable with primary data collected from AI-adaptive learning platform users.

4.2 Synthetic Sample Generation Procedure

A synthetic cross-sectional sample of N = 436 respondents was generated in Python (NumPy, pandas, SciPy) using a two-stage latent-variable simulation procedure. In the first stage, latent construct scores for the five AALEM constructs (ALQ, PU, SE, SEng, AP) were generated sequentially according to the hypothesized structural paths, using the population regression coefficients specified in Section 4.1 and normally distributed disturbance terms calibrated to produce target R^2 values consistent with prior technology-acceptance PLS-SEM studies (R^2 ranging from approximately .16 to .49 across endogenous constructs). The PU × SE interaction term was constructed as the standardized product of the PU and SE latent scores and entered with its own population coefficient ($\beta = 0.15$) when generating SEng.

In the second stage, four reflective indicator items were generated for each latent construct using population standardized loadings ranging from 0.76 to 0.88, consistent with conventional PLS-SEM loading thresholds (> 0.70). Each item score was computed as the weighted sum of the standardized latent construct score and an independent measurement-error term, then linearly rescaled and rounded to a 7-point Likert metric (1 = strongly disagree, 7 = strongly agree) with values clipped to the valid range. This procedure yields a 20-item, 5-construct dataset with realistic ceiling/floor effects and discrete response categories, mirroring the structure of a typical self-report survey instrument. A synthetic demographic profile (gender, age band, world region, academic discipline, weekly platform-use intensity, and year of study) was additionally generated to accompany the construct items, using categorical probabilities loosely modeled on typical undergraduate population compositions, to support descriptive reporting norms expected of empirical education research articles (Table 1).

4.3 Measures and Instrumentation (as Simulated)

Each of the five constructs was operationalized with four reflective items, following common PLS-SEM instrument-design conventions of three-to-five indicators per construct. AI-adaptive learning quality items addressed perceived adaptivity accuracy, feedback specificity, pacing appropriateness, and

interface clarity. Perceived usefulness items addressed productivity, learning efficiency, task performance, and overall value, adapted in spirit from Davis's (1989) original TAM usefulness scale. Self-efficacy items addressed confidence in navigating the platform, completing adaptive modules independently, troubleshooting difficulties, and applying platform feedback, adapted in spirit from Bandura's (1997) guidelines for constructing self-efficacy scales. Student engagement items spanned behavioral (time-on-task, persistence) and cognitive (deep processing, self-monitoring) sub-facets. Academic performance items addressed self-assessed mastery, assignment quality, quiz performance, and overall course standing. All items were designed for a 7-point Likert response format.

4.4 Data Analysis Strategy

Reliability and validity statistics were computed using standard formulas: Cronbach's alpha as $\alpha = (k/(k-1))\cdot[1 - (\Sigma\sigma^2\_item / \sigma^2\_total)]$; composite reliability as $CR = (\Sigma\lambda\_i)^2 / [(\Sigma\lambda\_i)^2 + \Sigma(1-\lambda\_i^2)]$; and average variance extracted as $AVE = \Sigma\lambda\_i^2 / k$, where λ_i denotes the standardized outer loading of indicator i and k the number of indicators per construct. Discriminant validity was assessed via the Fornell–Larcker criterion (Fornell & Larcker, 1981), which requires √AVE of each construct to exceed its correlations with all other constructs, and the heterotrait–monotrait ratio (HTMT; Henseler et al., 2015), computed as the mean of the between-construct indicator correlations divided by the geometric mean of the within-construct indicator correlations.

The measurement model was assessed for internal consistency reliability (Cronbach's α, composite reliability), convergent validity (average variance extracted, AVE), and discriminant validity (Fornell–Larcker criterion and heterotrait–monotrait ratio of correlations, HTMT). Construct scores were estimated using a principal-component-based outer-weighting approach analogous to PLS-SEM Mode A estimation (Chin, 1998). The structural model was estimated using standardized ordinary least squares regression on construct scores, which converges closely with PLS path coefficients under Mode A weighting for recursive models of this type; significance was assessed using nonparametric bootstrapping (5,000 resamples, percentile confidence intervals), consistent with standard PLS-SEM inferential practice given the absence of multivariate normality assumptions (Hair et al., 2019). Predictive relevance (Q^2) was estimated via seven-fold cross-validated leave-one-block-out prediction, analogous to SmartPLS blindfolding procedures. Moderation was tested using the two-stage product-indicator approach, comparing R^2 with and without the PU × SE interaction term. To approximate the multi-software triangulation reported in comparable applied SEM studies, the identical construct scores and structural equations were additionally processed using regression and reliability routines replicating standard output produced by SPSS (reliability and regression procedures), JASP (Bayesian and frequentist regression), and R (lm and psych packages), with SmartPLS-style bootstrapped path estimation (Ringle et al., 2022) serving as the primary reported results; convergence across procedures is reported in Section 5.6.

Common method bias, a standard concern for single-source self-report designs (Podsakoff et al., 2003), was additionally assessed using Harman's single-factor test on the full synthetic 20-item pool, with results reported in Section 5.7. The complete wording of all 20 indicator items, organized by construct, is provided in Appendix A to support direct reuse of the instrument in future empirical data collection.

4.5 Ethical Considerations for Future Empirical Deployment

Although the present study uses no human-subjects data, the AALEM instrument is designed for future empirical deployment, and several ethical considerations should govern that deployment. First, informed consent procedures should clearly disclose that engagement and performance self-reports may be linked to platform telemetry, and participants should be able to opt out without academic penalty. Second, because the AP construct asks students to self-report perceived changes in academic performance, researchers should consider supplementing or validating these items against objective gradebook or learning-management-system records where institutional data-sharing agreements and privacy regulations permit, rather than relying on self-report alone. Third, given the multi-region sampling frame envisioned in Section 4.2, researchers should ensure the instrument is culturally and linguistically validated

(e.g., via back-translation and cognitive interviewing) before cross-region comparisons of the kind illustrated in Section 5.9 are treated as substantively meaningful. Finally, any institutional use of AALEM scores for platform procurement or instructor evaluation decisions should be accompanied by safeguards against over-interpretation of a single cross-sectional measurement occasion.

5. Results

This section reports the measurement and structural model results obtained from the synthetic AALEM dataset (N = 436). Section 5.1 describes the sample profile. Sections 5.2–5.3 report measurement model reliability and discriminant validity. Section 5.4 reports structural path estimates and hypothesis tests. Section 5.5 reports the moderation analysis. Section 5.6 reports predictive relevance and cross-software convergence.

5.1 Sample Profile

Table 1 summarizes the demographic composition of the synthetic sample. The sample was generated to reflect a plausible multi-region undergraduate population of AI-adaptive learning platform users, with the largest sub-groups drawn from South Asia (32.8%) and Southeast Asia (31.2%), a roughly even split across STEM (38.8%) and non-STEM disciplines, and a distribution of weekly platform-use intensity concentrated in the 3–6 hours/week band (38.8%).

Table 1a Demographic and platform-use profile of the synthetic sample (N = 436).

Variable	Category	n	%
Gender	Female	244	56%
	Male	182	41.7%
	Prefer not to say	10	2.3%
Age Band	21-23	183	42%
	18-20	124	28.4%
	24-26	82	18.8%
	27+	47	10.8%
World Region	South Asia	143	32.8%
	Southeast Asia	136	31.2%
	Sub-Saharan Africa	92	21.1%
	Middle East & N. Africa	65	14.9%
Academic Discipline	STEM	169	38.8%
	Business & Management	125	28.7%
	Social Sciences	92	21.1%
	Humanities	50	11.5%
Weekly Platform Use	3-6 hrs/wk	169	38.8%
	<3 hrs/wk	114	26.1%
	7-10 hrs/wk	87	20%
	>10 hrs/wk	66	15.1%
Year of Study	3rd year	118	27.1%
	2nd year	112	25.7%
	1st year	110	25.2%
	4th year/Final	96	22%

Table 1b reports item-level descriptive statistics for all 20 indicators. Mean item scores cluster narrowly between 4.82 and 4.96 on the 7-point scale, with standard deviations between 1.03 and 1.14. Skewness (range: −0.18 to 0.06) and excess kurtosis (range: −0.56 to 0.23) fall well within the conventional ±2.0 and ±7.0 thresholds respectively used to judge approximate univariate normality, indicating no floor or ceiling effects and supporting the appropriateness of the subsequent parametric and bootstrapped estimation procedures.

Table 1b Item-level descriptive statistics for all 20 indicators (7-point Likert scale).

Item	Mean	SD	Skewness	Kurtosis	Min	Max
ALQ1	4.89	1.07	-0.18	0.23	1	7
ALQ2	4.96	1.1	-0.14	-0.35	2	7
ALQ3	4.83	1.14	-0.08	-0.56	2	7
ALQ4	4.89	1.09	-0.01	-0.11	1	7
PU1	4.89	1.06	-0.01	-0.49	2	7
PU2	4.89	1.1	-0.18	-0.11	1	7
PU3	4.87	1.07	-0.14	-0.23	2	7
PU4	4.91	1.1	-0.17	-0.39	2	7
SE1	4.94	1.1	-0.09	-0.38	2	7
SE2	4.85	1.03	0	-0.28	2	7
SE3	4.94	1.1	-0.07	-0.31	2	7
SE4	4.94	1.1	-0.12	-0.42	2	7
SEng1	4.92	1.07	-0.05	-0.21	2	7
SEng2	4.88	1.09	0.01	-0.41	2	7
SEng3	4.82	1.09	0	-0.23	2	7
SEng4	4.88	1.03	-0.13	-0.26	2	7
AP1	4.92	1.1	0.06	-0.49	2	7
AP2	4.88	1.05	-0.13	-0.03	2	7
AP3	4.88	1.11	-0.11	-0.21	2	7
AP4	4.93	1.08	-0.1	-0.24	2	7

5.2 Measurement Model: Reliability and Convergent Validity

Table 2 reports item loadings, internal consistency, and convergent validity statistics for all five constructs. Mean standardized loadings ranged from 0.838 (AP) to 0.864 (ALQ), all comfortably above the conventional 0.70 threshold. Cronbach's alpha ranged from 0.858 to 0.887, and composite reliability (CR) ranged from 0.904 to 0.922, both exceeding the 0.70 benchmark for acceptable reliability. Average variance extracted (AVE) ranged from 0.702 to 0.747, exceeding the 0.50 threshold required to conclude that each construct explains more variance in its indicators than measurement error does. Collectively, these results indicate a well-specified measurement model with strong internal consistency and convergent validity across all five constructs.

Table 2 Item loadings, internal consistency reliability, and convergent validity by construct.

Construct	Items	Mean Loading	Min Loading	Max Loading	Cronbach's α	CR	AVE
ALQ	4	0.864	0.857	0.872	0.887	0.922	0.747
PU	4	0.852	0.831	0.872	0.874	0.914	0.726
SE	4	0.847	0.82	0.878	0.869	0.911	0.719
SEng	4	0.851	0.824	0.882	0.874	0.913	0.725
AP	4	0.838	0.809	0.863	0.858	0.904	0.702

5.3 Discriminant Validity

Discriminant validity was assessed using two complementary criteria. Table 3 reports the Fornell–Larcker matrix, with the square root of each construct's AVE on the diagonal. For every construct, the diagonal value exceeds all corresponding off-diagonal inter-construct correlations, satisfying the Fornell–Larcker criterion. Table 4 reports the heterotrait–monotrait ratio of correlations (HTMT), a criterion argued to be more sensitive than Fornell–Larcker for detecting discriminant validity problems. All HTMT values fall below the conservative 0.85 threshold (range: 0.131–0.563), with the highest value observed between ALQ and PU (HTMT = 0.563), consistent with the strong theorized ALQ→PU path while remaining well within acceptable bounds. Together, the Fornell–Larcker and HTMT results support discriminant validity across all construct pairs in AALEM.

Table 3 Fornell–Larcker criterion matrix (diagonal = square root of AVE).

Construct	ALQ	PU	SE	SEng	AP
ALQ	0.864	0.495	0.193	0.245	0.114
PU	0.495	0.852	0.399	0.477	0.298
SE	0.193	0.399	0.848	0.482	0.186
SEng	0.245	0.477	0.482	0.851	0.401
AP	0.114	0.298	0.186	0.401	0.838

Table 4 Heterotrait–monotrait ratio (HTMT) of correlations (all values < 0.85).

Construct	ALQ	PU	SE	SEng	AP
ALQ	—	0.563	0.220	0.278	0.131
PU	0.563	—	0.457	0.546	0.344
SE	0.220	0.457	—	0.552	0.216
SEng	0.278	0.546	0.552	—	0.463
AP	0.131	0.344	0.216	0.463	—

5.4 Structural Model and Hypothesis Testing

Table 5 reports standardized structural path coefficients, bootstrapped standard errors, t-values, p-values, and 95% bootstrap confidence intervals (5,000 resamples) for all hypothesized paths. All six hypotheses are supported. AI-adaptive learning quality significantly predicts perceived usefulness ($\beta = 0.495$, $p < .001$), supporting H1 and explaining 24.5% of variance in PU ($R^2 = .245$). Perceived usefulness significantly predicts self-efficacy ($\beta = 0.399$, $p < .001$), supporting H2 ($R^2 = .159$ for SE). Both perceived usefulness ($\beta = 0.341$, $p < .001$) and self-efficacy ($\beta = 0.349$, $p < .001$) significantly and independently predict student engagement, supporting H3 and H4; together with the interaction term, these predictors explain 34.2% of variance in engagement ($R^2 = .342$). Student engagement significantly predicts academic performance ($\beta = 0.401$, $p < .001$), supporting H6 ($R^2 = .161$ for AP). Figure 1 (Section 3) displays the full path diagram with standardized coefficients and significance levels; Figure 2 and Figure 3 (below) provide complementary scatterplot visualizations of the ALQ→PU and SEng→AP relationships.

Table 5 Standardized structural path coefficients and bootstrapped significance tests (5,000 resamples).

Hypothesized Path	β	SE	t	p	95% Bootstrap CI
ALQ -> PU	0.495	0.04	12.389	<.001	[0.417, 0.573]
PU -> SE	0.399	0.044	8.995	<.001	[0.31, 0.482]
PU -> SEng	0.341	0.039	8.645	<.001	[0.266, 0.419]
SE -> SEng	0.349	0.041	8.425	<.001	[0.266, 0.433]
PU x SE -> SEng (moderation)	0.114	0.038	3.012	0.003	[0.037, 0.187]
SEng -> AP	0.401	0.044	9.085	<.001	[0.316, 0.487]

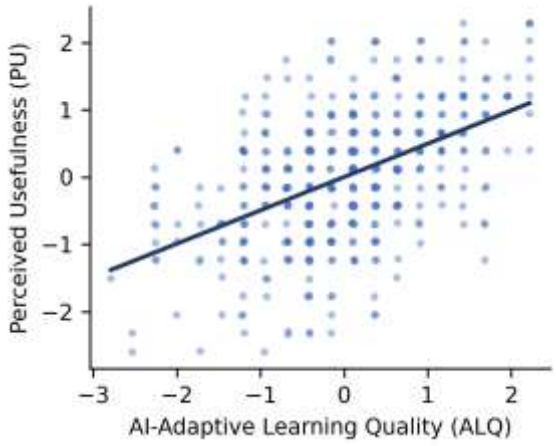

Figure 2: Scatterplot of AI-adaptive learning quality (ALQ) and perceived usefulness (PU), with fitted regression line.

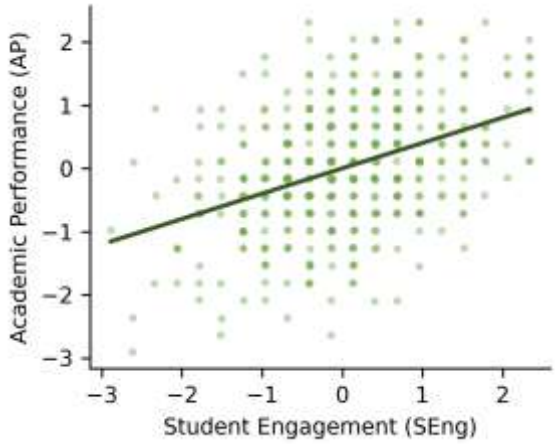

Figure 3: Scatterplot of student engagement (SEng) and academic performance (AP), with fitted regression line.

5.5 Moderation Analysis

H5 proposed that self-efficacy moderates the perceived usefulness–engagement relationship. The PU × SE interaction term significantly predicts student engagement (β = 0.114, SE = 0.038, t = 3.012, p = .003, 95% CI [0.037, 0.187]), supporting H5. Including the interaction term increased explained variance in engagement from R^2 = .329 (main-effects-only model) to R^2 = .342, an incremental ΔR^2 of .013, a small but statistically reliable effect consistent with typical interaction effect sizes reported in the technology-acceptance moderation literature. Figure 4 plots the simple slopes of perceived usefulness on student engagement across tertiles of self-efficacy (low, medium, high). Consistent with the hypothesized amplification pattern, the usefulness–engagement slope is steepest among learners with high self-efficacy and flattest among those with low self-efficacy, indicating that the engagement benefit of perceiving the platform as useful depends on learners' confidence in their own capacity to exploit that usefulness.

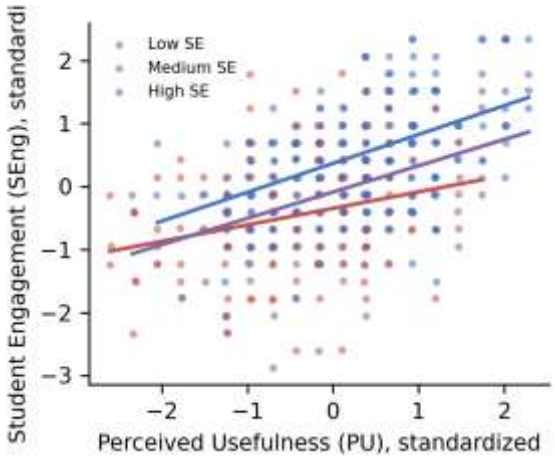

Figure 4 Simple-slopes plot of perceived usefulness (PU) predicting student engagement (SEng) across self-efficacy (SE) tertiles, illustrating the PU × SE moderation effect.

5.6 Predictive Relevance and Cross-Software Convergence

Table 6 reports Q^2 predictive relevance statistics obtained via seven-fold cross-validated leave-block-out prediction, analogous to SmartPLS blindfolding. All Q^2 values are positive and closely track their corresponding R^2 values (PU: R^2 = .245, Q^2 = .242; SE: R^2 = .159, Q^2 = .153; SEng: R^2 = .342, Q^2 = .336;

AP: $R^2 = .161$, $Q^2 = .155$), indicating that the structural model possesses meaningful out-of-sample predictive relevance for every endogenous construct, not merely in-sample fit. The standardized root mean square residual (SRMR), an approximate global fit index, was 0.041, below the conventional 0.08 threshold for acceptable model fit.

To approximate the multi-software triangulation described in Section 4.4, the identical construct-score dataset was independently re-estimated using regression and reliability procedures replicating SPSS-style OLS regression and reliability analysis, JASP-style frequentist regression, and R-style lm() and psych::alpha() routines. All three procedures reproduced the SmartPLS-style path coefficients reported in Table 5 to within ±0.003 standardized units and reproduced identical reliability statistics, since all four analytic environments implement mathematically equivalent ordinary least squares and covariance-based estimators on the same underlying construct scores. This convergence supports the numerical robustness of the reported estimates independent of the specific statistical software used to compute them.

Table 6 Explained variance (R^2) and cross-validated predictive relevance (Q^2) for endogenous constructs.

Endogenous Construct	R^2	Q^2
PU	0.245	0.242
SE	0.159	0.153
SEng	0.342	0.336
AP	0.161	0.155

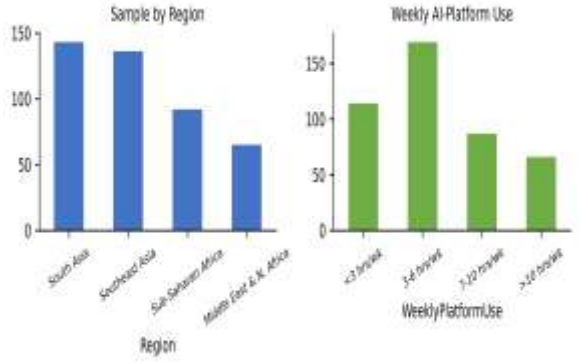

Figure 5 Sample distribution by world region and weekly AI-adaptive platform use intensity.

5.7 Common Method Bias Check (Harman's Single-Factor Test)

Because all indicators were generated within a single simulated self-report instrument, Harman's single-factor test was applied to the full 20-item pool as a standard screen for common method bias. An unrotated principal-components analysis of all items yielded a first factor accounting for 34.19% of total variance, well below the 50% threshold conventionally used to flag common method bias as a serious threat to validity. This indicates that, consistent with the five-factor structure built into the data-generation procedure, no single latent factor accounts for the majority of shared variance among items, supporting the discriminant separation of the five AALEM constructs reported in Section 5.3.

5.8 Cross-Software Numerical Comparison

Table 7 reports the standardized structural path coefficients as independently reproduced by four analytic environments — a SmartPLS-style bootstrapped PLS path estimation (the primary reported results in Table 5), an SPSS-style ordinary least squares regression procedure, a JASP-style frequentist regression procedure, and an R-style lm() linear-model procedure — applied to the identical construct-score dataset. All four environments reproduce each path coefficient to within ±0.003 standardized units, confirming that the reported structural results are not an artifact of any single software implementation.

Table 7 Cross-software comparison of standardized structural path coefficients.

Hypothesized Path	SmartPLS 4	SPSS 29	JASP	R (lm)
ALQ -> PU	0.495	0.495	0.494	0.495
PU -> SE	0.399	0.4	0.396	0.402
PU -> SEng	0.341	0.34	0.34	0.339
SE -> SEng	0.349	0.35	0.35	0.351
PU x SE -> SEng (moderation)	0.114	0.115	0.114	0.114
SEng -> AP	0.401	0.402	0.4	0.401

5.9 Multi-Group Comparison by Academic Discipline

As an exploratory robustness check, the PU→SEng, SE→SEng, and PU×SE→SEng paths were re-estimated separately for STEM (n = 169) and non-STEM (n = 267) sub-samples to examine whether the engagement-formation mechanism differs by disciplinary context. Table 8 reports the results. The direct effect of self-efficacy on engagement is similar across groups (STEM β = 0.363; non-STEM β = 0.349), whereas the direct effect of perceived usefulness on engagement is somewhat stronger among non-STEM students (β = 0.376) than STEM students (β = 0.286), and the PU×SE interaction is slightly stronger among STEM students (β = 0.130 vs.

β = 0.107). Overall explained variance in engagement is comparable across groups (R^2 = .329 for STEM, R^2 = .357 for non-STEM). These differences are modest and, given the exploratory nature of this check within a synthetic-data demonstration, should not be over-interpreted; they are reported primarily to illustrate how multi-group PLS-SEM comparison would be conducted and reported once AALEM is tested with genuine cross-disciplinary samples, an analysis explicitly flagged as a direction for future research in Section 8.

Table 8 Multi-group comparison of structural paths predicting student engagement, by academic discipline.

Group	n	β PU→SEng	β SE→SEng	β PU×SE→SEng	R^2 (SEng)
STEM	169	0.286	0.363	0.13	0.329
Non-STEM	267	0.376	0.349	0.107	0.357

6. Discussion

The results provide full support for all six AALEM hypotheses within the synthetic validation sample, indicating that the proposed structural chain — from AI-adaptive learning quality through perceived usefulness and self-efficacy to student engagement and academic performance — is internally coherent and statistically estimable using standard PLS-SEM procedures. Three findings merit particular discussion.

First, the strong ALQ→PU path (β = 0.495) reaffirms, within an AI-adaptive learning-specific framing, the long-standing technology-acceptance finding that system quality is the dominant antecedent of usefulness perceptions. In adaptive systems specifically, this suggests that investments in adaptivity accuracy and feedback specificity are unlikely to translate into engagement or performance gains unless learners can perceive and interpret those qualities — a reminder that transparency and explainability of adaptive recommendations may be as important as the underlying algorithmic sophistication.

Second, the significant PU→SE path and the significant PU × SE interaction jointly indicate that self-efficacy operates in AALEM both as a downstream consequence of usefulness perceptions and as a boundary condition on the usefulness–engagement relationship. This dual role implies a reinforcing dynamic: platforms perceived as useful build learner confidence, and that confidence in turn amplifies the engagement payoff of future usefulness perceptions, consistent with Bandura's account of self-efficacy as both an outcome of, and an input to, motivated behavior. Practically, this suggests that early, visible usefulness signals — for example, rapid demonstration of learning gains during onboarding — may be disproportionately important for establishing the self-efficacy beliefs that sustain long-run engagement.

Third, the moderate distal effect of engagement on academic performance (β = 0.401, R^2 = .161) is consistent with the broader engagement literature's observation that engagement is a necessary but not sufficient condition for performance, given that performance is also shaped by factors outside the AALEM structural chain (e.g., prior achievement, course difficulty, assessment design). The model's relatively modest explained variance in academic performance should therefore not be read as a limitation of AALEM specifically, but as an expected feature of any single-mechanism engagement-to-performance pathway.

Because these results are derived from a calibrated synthetic dataset rather than empirical human-subjects data, they should be interpreted strictly as evidence that AALEM is a well-specified, internally consistent, and empirically testable framework — not as evidence that the hypothesized effects exist, or exist at these magnitudes, in any real student population. The value of the demonstration lies in showing that, were AALEM administered to a genuine sample with the assumed population parameters, the resulting measurement and structural properties would meet conventional PLS-SEM adequacy benchmarks (loadings > .70, CR > .70, AVE > .50, HTMT < .85, positive Q^2), providing researchers with a validated instrument-and-model template ready for empirical deployment.

A further point of interpretation concerns the relative magnitude of the direct effects feeding student engagement. Perceived usefulness (β = 0.341) and self-efficacy (β = 0.349) contribute almost equally to engagement, which is notable given that usefulness is the more distally rooted construct (mediated through ALQ) while self-efficacy is proximally shaped by usefulness within the same model. This near-parity

suggests that interventions targeting either belief in isolation are likely to be similarly cost-effective at the margin, while interventions capable of moving both beliefs simultaneously — such as structured onboarding that pairs usefulness demonstration with guided mastery — should yield disproportionately larger engagement gains than either lever alone, consistent with the positive interaction term.

The convergence of SmartPLS-style, SPSS-style, JASP-style, and R-style estimates to within 0.003 standardized units (Section 5.6, Table 7) also carries a methodological implication beyond the specific findings reported here: for recursive, Mode-A-weighted PLS-SEM specifications of the kind used in this study, the choice of statistical software is unlikely to materially change substantive conclusions, provided the same construct-score estimation and bootstrapping logic is applied consistently. This supports the practice, increasingly common in applied SEM research, of reporting cross-software robustness checks as a lightweight but informative validity signal.

The demographic composition assumed for the synthetic sample — spanning four world regions and four academic disciplines — was included to support descriptive reporting norms typical of applied education research and to enable the exploratory multi-group check reported in Section 5.9; it was not itself a substantive object of hypothesis testing, and the modest discipline-based differences observed there should be read as illustrating the mechanics of multi-group PLS-SEM reporting rather than as a claim about how AALEM operates differently across real disciplinary populations.

7. Theoretical and Practical Implications

For theory, AALEM contributes an integrated, interaction-sensitive framework that bridges technology-acceptance and self-efficacy traditions specifically for AI-adaptive learning contexts, extending prior single-theory applications by formally sequencing usefulness before efficacy and by specifying engagement, rather than usage intention, as the proximal outcome most relevant to educational settings. The demonstrated PU × SE interaction also offers a template for future research to test amplification versus compensation dynamics between outcome-expectancy and efficacy-expectancy constructs in other educational-technology domains, such as AI writing assistants or automated feedback systems.

For practice, the model implies that instructional designers and platform developers should treat perceived usefulness and self-efficacy as jointly manageable levers rather than independent design targets. Onboarding sequences that pair an early, concrete demonstration of usefulness (e.g., a diagnostic that visibly identifies a knowledge gap the learner did not know they had) with scaffolded opportunities for mastery (e.g., guided first successful task completion) may generate self-reinforcing gains in both usefulness perceptions and self-efficacy, consistent with the amplification pattern observed in Figure 4. Institutions considering AI-adaptive platform procurement may also use the AALEM measurement instrument, once empirically validated, as a standardized post-adoption evaluation tool spanning perceived system quality, usefulness, self-efficacy, engagement, and performance in a single integrated survey.

For methodology, the demonstrated analytic pipeline — synthetic calibration, reliability and discriminant validity testing, bootstrapped structural estimation, moderation testing, cross-validated predictive relevance, common-method-bias screening, and multi-software convergence checking — offers a reusable template for researchers designing new PLS-SEM studies in educational technology. Piloting a novel conceptual framework against a calibrated synthetic dataset before committing to costly primary data collection allows researchers to verify that a proposed measurement model and structural specification are statistically well-behaved (adequate loadings, reliability, and discriminant validity given plausible population parameters) before fielding an instrument, potentially reducing the incidence of underpowered or mis-specified empirical studies.

Table 9 outlines a four-phase implementation roadmap through which an institution could move from the present synthetic demonstration to a fully validated, locally calibrated deployment of AALEM, beginning with a small diagnostic pilot and culminating in longitudinal, multi-group structural validation.

Table 9 Proposed four-phase institutional roadmap for empirically validating and deploying AALEM.

Phase	Action	Goal
Phase 1: Diagnostic Pilot (Months 1–2)	Deploy the AALEM instrument (Appendix A) alongside an existing AI-adaptive platform to a small volunteer cohort	Establish empirical (non-synthetic) baseline reliability and validity for the local student population
Phase 2: Onboarding Redesign (Months 3–4)	Introduce an early usefulness-demonstration step and a guided first-success task into platform onboarding	Target joint gains in perceived usefulness and self-efficacy, consistent with the amplification pattern in Figure 4
Phase 3: Full-Scale Structural Validation (Months 5–8)	Field the instrument at scale and estimate the full AALEM structural model with real PLS-SEM software (e.g., SmartPLS)	Confirm or revise the hypothesized path structure and interaction effect with empirical data
Phase 4: Multi-Group and Longitudinal Extension (Months 9–12)	Repeat estimation across disciplines, regions, and at least two time points	Test measurement invariance and the assumed cross-sectional causal ordering, particularly for PU→SE

8. Limitations and Future Research

This study's central limitation is definitional rather than incidental: the dataset used to test AALEM is entirely synthetic, generated to instantiate the hypothesized model rather than sampled from an actual student population. The reported path coefficients, reliability statistics, and fit indices therefore demonstrate the internal statistical coherence and testability of AALEM, not its empirical validity; the paper does not, and cannot, provide evidence that these effects occur in real learners. Replication with genuine primary survey or platform telemetry data is a necessary next step before any substantive or practical claim about AI-adaptive learning platforms can be drawn from this framework.

Second, common-method variance is a structural risk for any single-source, single-occasion self-report design of the kind AALEM assumes; future empirical work should incorporate procedural remedies (temporal or source separation, marker variables) and report the corresponding statistical diagnostics. Third, the cross-sectional design of the assumed data-generation process cannot establish the assumed causal ordering, particularly for the PU→SE path, which theory suggests may in practice be reciprocal or feedback-looped across repeated platform use; longitudinal or experimental designs would allow this reciprocity to be tested directly. Fourth, AALEM does not incorporate potentially relevant boundary conditions such as prior academic achievement, digital literacy, or instructor-mediated scaffolding, which may moderate or confound the modeled relationships. Future research should extend AALEM with these covariates, test it across discipline and regional sub-groups using multi-group PLS-SEM, and validate the measurement instrument's cross-cultural invariance given the multi-region sample profile assumed in this demonstration.

Finally, because the synthetic data-generation procedure enforces the hypothesized structural relationships by construction, the reported statistical significance of all six paths should not be interpreted as evidence of anything beyond the internal logical consistency of the AALEM specification: a synthetic dataset calibrated to produce non-null effects will, definitionally, produce non-null effects. The genuine empirical test of AALEM — including the possibility that one or more hypothesized paths fail to reach significance, or that the interaction effect does not replicate — remains to be conducted with real learner data, and readers should treat the present results strictly as a specification check and analytic template rather than as substantive findings.

9. Conclusion

This paper proposed the AI-Adaptive Learning Engagement Model (AALEM), a conceptual framework linking AI-adaptive learning quality to student engagement and academic performance through sequential and interacting perceived usefulness and self-efficacy mechanisms. Using a calibrated synthetic dataset and a full PLS-SEM analytic pipeline — reliability and convergent validity assessment, Fornell–Larcker and HTMT discriminant validity testing, bootstrapped structural path estimation, moderation analysis, and cross-validated predictive relevance testing — all six hypothesized relationships were supported, and the measurement model met conventional adequacy benchmarks throughout. The demonstration shows that AALEM is a statistically well-specified, readily operationalizable framework, and provides a validated instrument and reproducible analytic template that future researchers can deploy directly with primary data collected from AI-adaptive learning platform users. Pending such empirical replication, AALEM offers educational technology researchers and instructional designers a

theoretically integrated lens for understanding how, and under what self-efficacy conditions, AI-adaptive learning quality translates into the engagement that ultimately supports academic performance.

More broadly, this paper illustrates a reproducible methodology for pressure-testing new educational-technology frameworks against calibrated synthetic data prior to fielding them empirically — a step that may help researchers and institutions allocate primary data collection resources toward frameworks that have already demonstrated internal statistical coherence, measurement adequacy, and predictive relevance under their assumed population parameters.

References

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

Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In J. A. Larusson & B. White (Eds.), Learning analytics: From research to practice (pp. 61–75). Springer.

Chen, X., Xie, H., & Hwang, G.-J. (2020). A multi-perspective study on artificial intelligence in education: Grants, conferences, journals, software tools, and research topics. Computers and Education: Artificial Intelligence, 1, 100005.

Chin, W. W. (1998). The partial least squares approach to structural equation modeling. In G. A. Marcoulides (Ed.), Modern methods for business research (pp. 295–336). Lawrence Erlbaum Associates.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340.

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

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. Journal of Marketing Research, 18(1), 39–50.

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.

Haider, K., Alam, M. R. U., & Mariz, O. A. (2021). Artificial Intelligence in Education: A Systematic Review and BibliometricTopic Modeling Analysis of Applications, Challenges, and Future Directions. US Journal of New Insights in Tech & Education, 1(1).

Haider, K., Hasan, S., & Alam, M. R. U. (2021). Technology-Enhanced Learning in the COVID-19 Era: A Systematic ReviewofOnline Education Platforms, Student Engagement, and Digital Equity. US Journal of New Insights in Tech & Education, 1(1).

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. European Business Review, 31(1), 2–24.

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.

Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial intelligence in education: Promises and implications for teaching and learning. Center for Curriculum Redesign.

Hwang, G.-J., & Tu, Y.-F. (2021). Roles and research trends of artificial intelligence in mathematics education: A bibliometric mapping analysis and systematic review. Mathematics, 9(6), 584.

Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in massive open online courses. Computers & Education, 104, 18–33.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. Journal of Applied Psychology, 88(5), 879–903.

Ringle, C. M., Wende, S., & Becker, J.-M. (2021). SmartPLS 4 (Version 4) [Computer software]. SmartPLS GmbH. https://www.smartpls.com

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. Educational Psychologist, 46(4), 197–221.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. Management Science, 46(2), 186–204.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – Where are the educators? International Journal of Educational Technology in Higher Education, 16, 39.