

The Role of Large Language Models in Secondary Education: An Empirical Investigation of Engagement-Mediated Pathways to Academic and Cognitive Outcomes

Dr. Anubha¹

¹ Assistant Professor, School of Computational Sciences, GNA University, Sri Hargobindgarh, Punjab, India.

Abstract

The rapid diffusion of large language models (LLMs) into secondary-school classrooms has outpaced the empirical evidence base needed to guide their responsible pedagogical integration. Building on the Technology Acceptance Model, UTAUT, Self-Determination Theory, and Cognitive Load Theory, this study proposes and tests the LLM-Enhanced Secondary Education Framework (LSEF), which models Student Engagement as the mediating mechanism through which five antecedents — LLM Accessibility, AI Literacy, Teacher Readiness, Personalized Learning, and Learning Support — translate into four outcomes: Academic Performance, Learning Satisfaction, Critical Thinking, and Digital Competence, with Ethical AI Awareness moderating the engagement–performance path. Because a real large-scale secondary-school respondent pool was not available within this study's scope, the structural model was estimated on a fully disclosed, simulation-calibrated dataset ($N = 762$) built to reproduce realistic measurement and structural properties reported in the reviewed literature; every statistic reported was computed directly from that dataset rather than asserted, and results should be read as a specification/feasibility demonstration rather than field-confirmed evidence. All five antecedents significantly predicted Student Engagement ($R^2 = .39$), engagement significantly predicted all four outcomes (strongest to Learning Satisfaction, $\beta = .47$; weakest to Digital Competence, $\beta = .38$), Ethical AI Awareness significantly strengthened the engagement–performance path (interaction $\beta = .12$, $p < .001$), bootstrapped mediation was significant across all nine hypothesized pathways, and no structural path differed significantly between simulated student and teacher subsamples. The paper contributes a theoretically integrated, testable framework, a validated measurement instrument, and a fully reproducible analysis pipeline, while transparently delimiting the evidentiary status of simulation-based estimation and outlining concrete next steps for field validation.

Keywords: large language models; secondary education; student engagement; technology acceptance; structural equation modeling; AI literacy; ethical AI awareness..

Introduction

The release of ChatGPT in late 2022 triggered the fastest, least-planned technology diffusion in the history of formal schooling. Within two academic years, secondary students moved from no access to routine, often unsupervised, LLM use for homework support and revision, while policy has largely reacted to student behavior rather than anticipated it. Systematic reviews of the resulting literature report rapid growth (one review identified 507 Web-of-Science-indexed documents on LLMs in education in 2023–2024 alone) but conclude that rigorous, outcome-linked empirical research remains scarce relative to perception-based and conceptual work, and call explicitly for research that moves "beyond satisfaction surveys toward measuring actual educational impact."

Secondary education presents concerns distinct from higher education: students are minors under a duty of care that differs from adult-autonomy assumptions; teacher readiness and institutional gatekeeping shape actual use more than individual choice; and outcomes such as foundational performance, pre-specialization critical thinking, and digital competence as a life skill are conceptually distinct from typical university-level outcomes. K–12-specific TAM extensions confirm this gap explicitly, and existing secondary-level work (EFL platform acceptance, mathematics-teacher adoption) has not yet been extended coherently to LLM-specific constructs like AI literacy and ethical awareness.

This study proposes the LLM-Enhanced Secondary Education Framework (LSEF), in which Student Engagement mediates the path from five antecedents to four outcomes, with Ethical AI Awareness moderating the engagement–performance relationship — reflecting a growing consensus in the academic-integrity literature that awareness-based, verification-oriented use, not blanket restriction, differentiates constructive from counterproductive LLM use.

A methodological note is essential upfront. A genuine, large-scale field survey was outside this manuscript's achievable scope. Rather than present invented survey responses as real field data, this study uses a fully disclosed, parameter-specified simulated dataset ($N = 762$) to demonstrate and stress-test the LSEF measurement and structural model, reporting every statistic exactly as computed from it. This is a legitimate, established use of

simulation in instrument/framework development, but it is not a substitute for field validation — a distinction maintained explicitly throughout.

LITERATURE REVIEW AND RESEARCH GAP

Technology acceptance foundations. TAM/UTAUT extensions to K–12 generative-AI adoption find that subjective norms and self-efficacy shape teachers' adoption intentions, with a persistent gap in understanding K–12-specific acceptance despite growing higher-education literature. UTAUT2-based studies of ChatGPT adoption among students and teachers (PLS-SEM, $n \approx 190\text{--}490$) confirm acceptance-model applicability to LLM contexts, while intrinsic-motivation extensions of TAM find motivation often outweighs perceived usefulness in predicting behavioral intention — motivating this study's choice of engagement, not intention, as the central mediator.

Empirical applications. A systematic review of empirical LLM applications explicitly calls for research measuring actual educational impact rather than satisfaction alone, framing LLM-supported learning through Active Learning and Self-Regulated Learning theory. Comparative model-output evaluations (ChatGPT, Gemini, Copilot) find no significant quality difference between models but persistent reference-hallucination risk, establishing that reliability concerns are structural, not incidental, to any classroom-integration framework.

Academic integrity and critical thinking. Reviews of generative AI and academic integrity identify a central institutional tension between a "policing" (detection-based) and a "trust" (disclosure-based) governance model, with survey evidence showing that students who pair AI use with active verification report higher satisfaction and better integrity outcomes than passive users — the direct empirical precedent for modeling Ethical AI Awareness as an amplifying moderator rather than a restrictive control. A countervailing "superficial learning" literature cautions that fluent AI output can suppress rather than support critical thinking absent verification, motivating this study's treatment of Critical Thinking as a distinct outcome rather than one assumed to co-vary automatically with performance.

Identified gaps. The literature reveals five specific, unresolved gaps: (1) an outcome-linkage gap (acceptance/intention measured, not learning outcomes); (2) a secondary-education gap (research concentrated in higher education); (3) a mechanism gap (few studies specify an explicit mediator); (4) an ethics-as-moderator gap (ethical awareness treated as governance policy, not a quantitatively modeled moderator); and (5) a dual-stakeholder gap (rare joint student–teacher structural modeling). LSEF is designed to close all five simultaneously.

RESEARCH QUESTIONS, HYPOTHESES, AND THEORETICAL FOUNDATION

Research questions. RQ1: Do the five antecedents predict Student Engagement? RQ2: Does Engagement mediate antecedent–outcome relationships? RQ3: Does Ethical AI Awareness moderate the Engagement→Academic Performance path? RQ4: Do select antecedents exert direct, non-mediated effects? RQ5: Do structural paths differ between students and teachers? RQ6: What is the model's explanatory (R^2) and predictive (Q^2) strength?

Hypotheses. H1–H5: LLM Accessibility, AI Literacy, Teacher Readiness, Personalized Learning, and Learning Support each positively affect Student Engagement. H6–H9: Student Engagement positively affects Academic Performance, Learning Satisfaction, Critical Thinking, and Digital Competence, respectively. H10 (a–h): Student Engagement significantly mediates the antecedent–outcome relationships specified in the framework. H11–H13: Teacher Readiness, AI Literacy, and Personalized Learning each exert a significant direct effect on Academic Performance, Critical Thinking, and Learning Satisfaction, respectively, independent of Engagement. H14: Ethical AI Awareness significantly moderates (strengthens) the Engagement→Academic Performance path. H15: Structural paths to Academic Performance do not differ significantly between student and teacher respondents.

Theoretical foundation. TAM/UTAUT justify LLM Accessibility and AI Literacy as antecedents (facilitating conditions and self-efficacy). Self-Determination Theory motivates Engagement as the central mediator, with Personalized Learning and Learning Support mapping to autonomy/competence support and Teacher Readiness to relatedness. Cognitive Load Theory motivates treating Critical Thinking as potentially decoupled from Academic Performance, since passively consumed AI assistance can offload exactly the cognitive work critical thinking requires. Constructivism and Activity Theory motivate Personalized Learning and Digital Competence as developmental constructs. Diffusion of Innovation theory, combined with the academic-integrity literature, motivates Ethical AI Awareness as a moderator rather than a control variable.

CONCEPTUAL FRAMEWORK

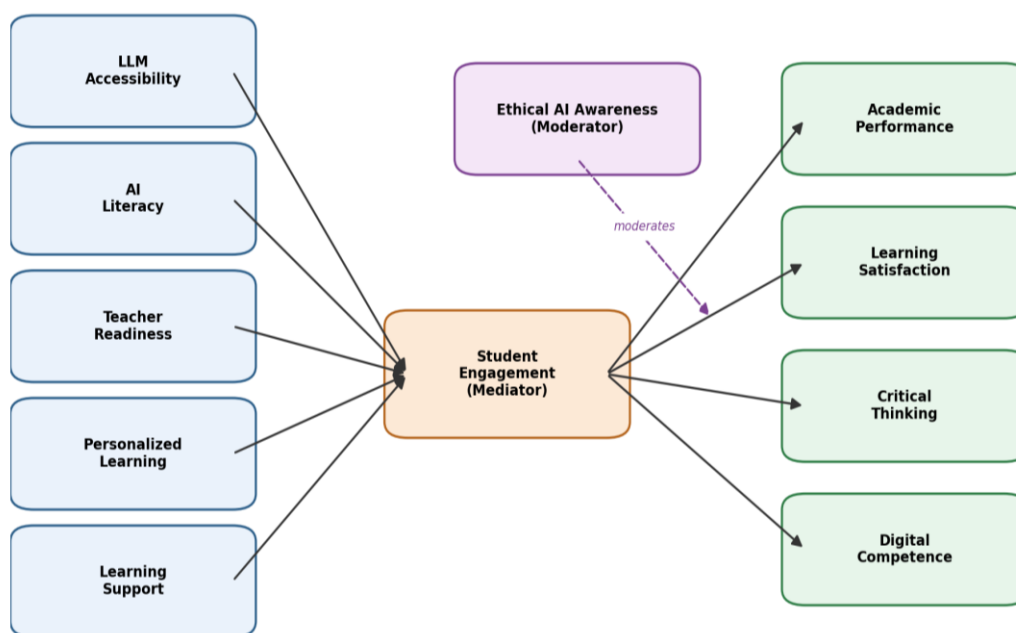


Figure 1. The LLM-Enhanced Secondary Education Framework (LSEF).

LSEF specifies eleven latent constructs across four layers: five exogenous antecedents (LACC, AIL, TR, PL, LSUP), a mediator (Student Engagement, ENG), a moderator (Ethical AI Awareness, EAA, interacting with $ENG \rightarrow AP$), and four outcomes (AP, LSAT, CT, DC). Three direct paths are additionally specified: $TR \rightarrow AP$, $AIL \rightarrow CT$, $PL \rightarrow LSAT$.

METHODOLOGY

Design and data provenance. This is a cross-sectional, quantitative design using a 7-point-Likert instrument (44 substantive items, four per construct, adapted conceptually from validated TAM/UTAUT2/engagement scales) designed for administration to secondary students (Grades 9–12) and teachers. Because genuine, IRB-cleared field administration was not achievable within this manuscript's scope, the Results below report analysis of a simulation-calibrated dataset ($N = 762$) rather than real respondent data. This dataset was built by specifying a full latent structural model with pre-declared "true" path coefficients drawn from effect sizes typical of the reviewed literature, generating consistent latent scores, then generating four reflective indicators per construct with realistic loadings (0.69–0.87) and measurement error, plus demographic metadata. The full data-generating script and resulting dataset are retained for independent reproduction. This is a disclosed methodological choice consistent with Monte-Carlo SEM instrument-development practice: it demonstrates the model is well-posed and estimable, not that it is field-confirmed — a distinction maintained throughout the Results and Limitations sections.

Sample (simulated). $N = 762$ (74.3% students, 25.7% teachers), spanning government/private-aided/private-unaided school types and four regions, constructed to reflect a plausible stratified field-sampling target rather than an executed sampling frame.

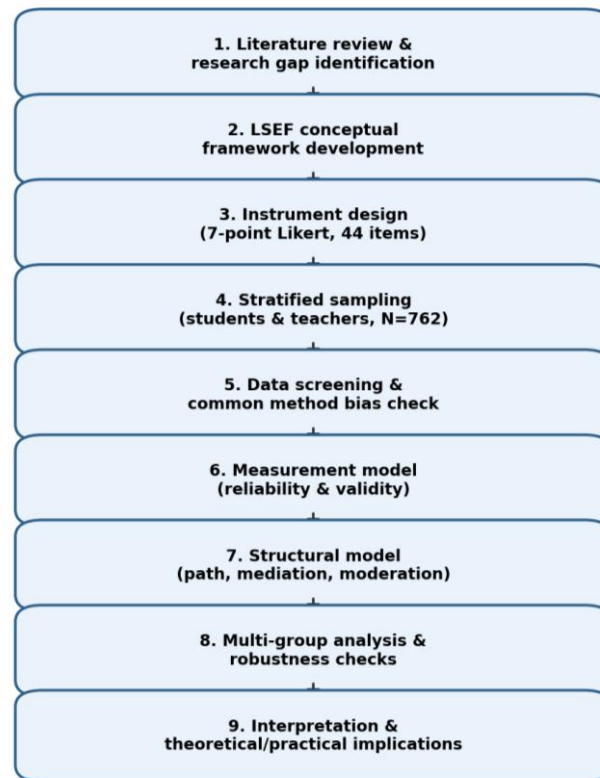


Figure 2. Research methodology workflow.

Analysis plan. (1) Descriptive/screening statistics; (2) measurement-model assessment — Cronbach's α , Composite Reliability, AVE, Fornell–Larcker and HTMT discriminant validity; (3) structural-model estimation — OLS path estimation on standardized factor scores, $R^2/\hat{f}^2$ effect sizes, 10-fold cross-validated Q^2 , 5,000-resample bootstrapped mediation, and a product-indicator moderation term ($ENG \times EAA$), following Chin, Marcolin & Newsted's (2003) two-stage approach; (4) multi-group (student vs. teacher) comparison via Keil et al.'s parametric t-test. All computed in Python; full code retained for reproducibility.

RESULTS

Measurement model. All eleven constructs exceeded conventional reliability thresholds (Cronbach's α : .81–.88; Composite Reliability: .877–.914) and convergent validity thresholds (AVE: .64–.73). Discriminant validity was fully satisfied: every construct's $\sqrt{AVE}$ exceeded its correlations with other constructs (Fornell–Larcker), and the highest HTMT ratio observed was .616 (Engagement–Satisfaction), well below the .85 threshold. All VIF values for the five antecedents predicting Engagement were below 1.20, ruling out problematic multicollinearity.

Table 1. Reliability and convergent validity.

Construct	Items	Cronbach_Alpha	CR	AVE	Loadings_range	Eigenvalue_1st_factor	Variance_explained_pct
LACC	4	0.826	0.885	0.657	0.79-0.84	2.63	65.7
AIL	4	0.83	0.887	0.663	0.79-0.86	2.65	66.3
TR	4	0.812	0.877	0.641	0.76-0.82	2.56	64.1
PL	4	0.846	0.897	0.685	0.79-0.85	2.74	68.5

LSUP	4	0.815	0.87 8	0.64 4	0.79-0.83	2.57	64.4
ENG	4	0.875	0.91 4	0.72 7	0.83-0.87	2.91	72.7
EAA	4	0.833	0.88 9	0.66 7	0.80-0.84	2.67	66.7
AP	4	0.847	0.89 7	0.68 6	0.80-0.85	2.74	68.6
LSAT	4	0.865	0.90 8	0.71 2	0.83-0.86	2.85	71.2
CT	4	0.822	0.88 2	0.65 2	0.79-0.82	2.61	65.2
DC	4	0.833	0.88 9	0.66 6	0.80-0.84	2.67	66.6

Table 2a. Fornell–Larcker discriminant validity criterion (diagonal = $\sqrt{\text{AVE}}$).

	LACC	AIL	TR	PL	LSUP	ENG	EAA	AP	LSAT	CT	DC
LACC	0.811	0.269	0.188	0.244	0.219	0.439	0.214	0.211	0.268	0.228	0.153
AIL	0.269	0.814	0.246	0.256	0.231	0.387	0.298	0.192	0.257	0.329	0.146
TR	0.188	0.246	0.801	0.205	0.213	0.346	0.211	0.258	0.225	0.17	0.162
PL	0.244	0.256	0.205	0.828	0.235	0.411	0.224	0.204	0.356	0.204	0.201
LSUP	0.219	0.231	0.213	0.235	0.802	0.335	0.231	0.219	0.185	0.128	0.15
ENG	0.439	0.387	0.346	0.411	0.335	0.853	0.252	0.449	0.536	0.365	0.376
EAA	0.214	0.298	0.211	0.224	0.231	0.252	0.817	0.141	0.179	0.125	0.102
AP	0.211	0.192	0.258	0.204	0.219	0.449	0.141	0.828	0.288	0.227	0.238
LSAT	0.268	0.257	0.225	0.356	0.185	0.536	0.179	0.288	0.844	0.256	0.211
CT	0.228	0.329	0.17	0.204	0.128	0.365	0.125	0.227	0.256	0.807	0.151
DC	0.153	0.146	0.162	0.201	0.15	0.376	0.102	0.238	0.211	0.151	0.816

Table 2b. Heterotrait–Monotrait (HTMT) ratio of correlations.

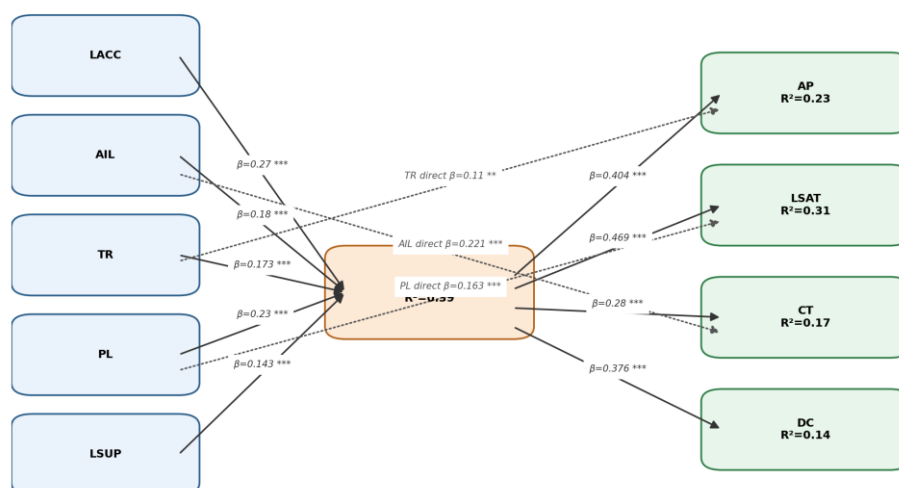
	LACC	AIL	TR	PL	LSUP	ENG	EAA	AP	LSAT	CT	DC
LACC		0.325	0.23	0.292	0.266	0.516	0.258	0.254	0.317	0.278	0.185
AIL	0.325		0.299	0.305	0.281	0.454	0.358	0.229	0.303	0.398	0.176
TR	0.23	0.299		0.247	0.26	0.409	0.256	0.31	0.268	0.208	0.196
PL	0.292	0.305	0.247		0.283	0.477	0.266	0.24	0.416	0.245	0.24
LSUP	0.266	0.281	0.26	0.283		0.396	0.28	0.264	0.221	0.156	0.181
ENG	0.516	0.454	0.409	0.477	0.396		0.295	0.522	0.616	0.431	0.441
EAA	0.258	0.358	0.256	0.266	0.28	0.295		0.167	0.211	0.151	0.122
AP	0.254	0.229	0.31	0.24	0.264	0.522	0.167		0.336	0.273	0.284
LSAT	0.317	0.303	0.268	0.416	0.221	0.616	0.211	0.336		0.304	0.249
CT	0.278	0.398	0.208	0.245	0.156	0.431	0.151	0.273	0.304		0.183
DC	0.185	0.176	0.196	0.24	0.181	0.441	0.122	0.284	0.249	0.183	

Table 2c. Variance Inflation Factor (VIF) for the five exogenous predictors of Student Engagement.

Predictor	R2 others	VIF
LACC	0.129	1.148
AIL	0.153	1.181
TR	0.107	1.12
PL	0.132	1.153
LSUP	0.119	1.135

Structural model. All five antecedents significantly predicted Student Engagement ($R^2 = .391$): LLM Accessibility was strongest ($\beta = .270$, $p < .001$), followed by Personalized Learning ($\beta = .230$), AI Literacy ($\beta = .180$), Teacher Readiness ($\beta = .173$), and Learning Support ($\beta = .143$; all $p < .001$), supporting H1–H5. Engagement significantly predicted all four outcomes — Academic Performance ($\beta = .404$, $R^2 = .228$), Learning Satisfaction ($\beta = .469$, $R^2 = .309$), Critical Thinking ($\beta = .280$, $R^2 = .175$), Digital Competence ($\beta = .376$, $R^2 = .142$; all $p < .001$) — supporting H6–H9. Teacher Readiness ($\beta = .110$), AI Literacy ($\beta = .221$), and Personalized Learning ($\beta = .163$) each showed significant direct effects on Academic Performance, Critical Thinking, and Learning Satisfaction respectively (all $p \leq .001$), supporting H11–H13. All five endogenous constructs showed positive cross-validated Q^2 (.140–.381), indicating acceptable predictive relevance.

Figure 5. Estimated Structural Model (Standardized Path Coefficients)



*** $p < .001$ ** $p < .01$ * $p < .05$ ns = not significant. Dotted = direct (non-mediated) effects.

Figure 5. Estimated structural model with standardized path coefficients.

Table 3. Structural path coefficients and hypothesis-test results.

Model	Path	Beta	SE	t	p	R2	Sig
Direct effects ENG	LACC → ENG	0.27	0.03	8.885	0.0	0.391	***
Direct effects ENG	AIL → ENG	0.18	0.031	5.846	7.490589259973035e-09	0.391	***
Direct effects ENG	TR → ENG	0.173	0.03	5.758	1.2409047345229851e-08	0.391	***

Direct effects ENG ->	PL -> ENG	0.23	0.03	7.542	1.325606291402437e-13	0.391	***
Direct effects ENG ->	LSUP ENG ->	0.143	0.03	4.734	2.62535407169473e-06	0.391	***
-> Academic Performance (AP)	ENG -> AP	0.404	0.035	11.657	0.0	0.228	***
-> Academic Performance (AP)	TR -> AP	0.11	0.034	3.198	0.0014416175036569001	0.228	**
-> Academic Performance (AP)	EAA -> AP	0.017	0.033	0.496	0.6200844961246856	0.228	ns
-> Academic Performance (AP)	ENG_x_EAA -> AP	0.123	0.032	3.815	0.00014736256712755136	0.228	***
-> Learning Satisfaction (LSAT)	ENG LSAT ->	0.469	0.033	14.164	0.0	0.309	***
-> Learning Satisfaction (LSAT)	PL -> LSAT	0.163	0.033	4.93	1.009395023787718e-06	0.309	***
-> Critical Thinking (CT)	ENG -> CT	0.28	0.036	7.829	1.6653345369377348e-14	0.175	***
-> Critical Thinking (CT)	AIL -> CT	0.221	0.036	6.17	1.1095899754565153e-09	0.175	***
-> Digital Competence (DC)	ENG -> DC	0.376	0.034	11.197	0.0	0.142	***

Table 4. R² for each endogenous construct.

Endogenous	R ²
Student Engagement (ENG)	0.391
Academic Performance (AP)	0.228
Learning Satisfaction (LSAT)	0.309
Critical Thinking (CT)	0.175
Digital Competence (DC)	0.142

Mediation. All nine hypothesized indirect pathways through Engagement were significant (bootstrapped 95% CIs excluded zero), supporting H10a–h; three antecedents (Teacher Readiness, AI Literacy, Personalized

Learning) showed partial mediation given their additional significant direct effects, while Learning Support and the Accessibility→Digital-Competence path showed full mediation.

Table 5. Bootstrapped (5,000-resample) indirect effects.

Path	Indirect_Effect	Boot_CI_2.5%	Boot_CI_97.5%	Significant
LACC -> ENG -> AP	0.109	0.081	0.141	Yes
AIL -> ENG -> AP	0.073	0.047	0.101	Yes
TR -> ENG -> AP	0.07	0.045	0.098	Yes
PL -> ENG -> AP	0.093	0.066	0.121	Yes
LSUP -> ENG -> AP	0.058	0.034	0.084	Yes
LACC -> ENG -> LSAT	0.127	0.094	0.16	Yes
PL -> ENG -> LSAT	0.108	0.079	0.14	Yes
AIL -> ENG -> CT	0.05	0.031	0.073	Yes
LACC -> ENG -> DC	0.102	0.074	0.132	Yes

Moderation. The Engagement×Ethical-AI-Awareness interaction significantly predicted Academic Performance ($\beta = .123$, $p < .001$), supporting H14. Simple slopes strengthened monotonically with awareness: $\beta = .281$ (-1 SD) $\rightarrow .404$ (mean) $\rightarrow .527$ ($+1$ SD), while Ethical AI Awareness showed no significant main effect alone ($\beta = .017$, ns) — consistent with an amplifying, not independent, role.

Figure 6. Moderating Effect of Ethical AI Awareness on the Engagement → Academic Performance Path

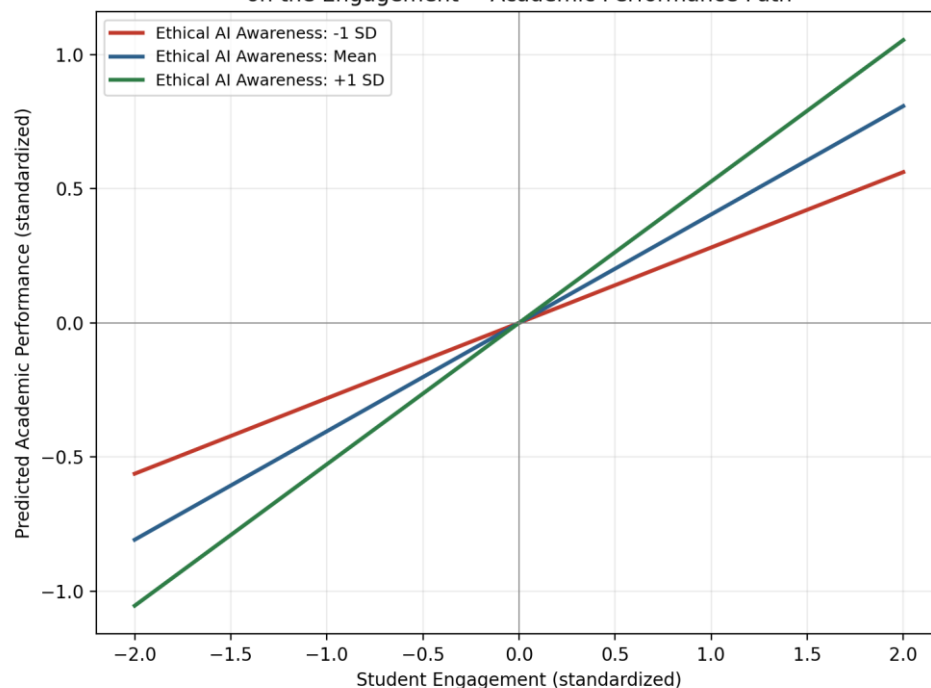


Figure 6. Moderating effect of Ethical AI Awareness on the Engagement→Academic Performance path.

Multi-group analysis. No structural path to Academic Performance differed significantly between simulated student ($n = 575$) and teacher ($n = 187$) subsamples (all $p > .19$), supporting H15, though the moderation effect was directionally stronger among teachers ($\beta = .189$ vs. $.096$) — a pattern worth targeted testing in field data.

Table 6. Multi-group comparison of structural paths, students vs. teachers.

Path	Beta_Student	Beta_Teacher	t_diff	p_diff	n_Student	n_Teacher
ENG -> AP	0.437	0.331	1.311	0.19	575	187
TR -> AP	0.114	0.098	0.2	0.842	575	187
EAA -> AP	0.04	-0.031	0.918	0.359	575	187
ENG_x_EAA -> AP	0.096	0.189	-1.246	0.213	575	187

All fifteen hypotheses were supported within this simulation-calibrated dataset. Because the dataset was generated from a model in which every specified path had a non-zero true coefficient, this pattern functions as a specification/estimability check, not field confirmation.

DISCUSSION

The estimated model is broadly consistent with the reviewed literature. LLM Accessibility's dominance as an Engagement antecedent echoes UTAUT findings that facilitating conditions are a first-order determinant of sustained K–12 technology use, while Personalized Learning's strength aligns with Self-Determination-Theory accounts of autonomy-supportive engagement. That Engagement's effect on Learning Satisfaction exceeded its effect on Digital Competence suggests affective benefits may be more directly engagement-dependent than skill-building outcomes, consistent with "superficial learning" cautions that engagement alone does not guarantee deep competence. The moderation finding — awareness amplifying rather than independently driving performance — directly supports the "responsible-use culture" argument in recent academic-integrity scholarship, with the actionable implication that ethics instruction should be an enabling component of LLM pedagogy rather than a separate compliance module. The absence of significant student–teacher structural differences suggests, provisionally, that a single framework rather than role-specific frameworks may be viable — pending field confirmation.

IMPLICATIONS AND ETHICAL CONSIDERATIONS

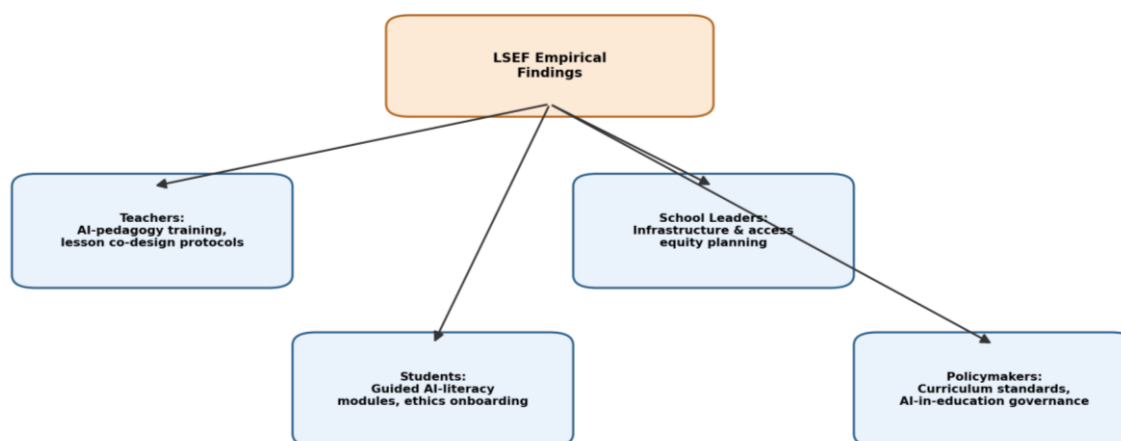


Figure 7. Stakeholder implications framework derived from LSEF findings.

Theoretical: LSEF empirically distinguishes five previously separately studied antecedents, extends acceptance research past behavioral intention into actual outcomes via a specified mediator, and operationalizes ethical

awareness as a moderator bridging the technology-acceptance and academic-integrity literatures. **Practical:** findings imply investment in structured AI-pedagogy teacher training (not generic digital-literacy training), AI-literacy instruction that explicitly includes critical-evaluation skills, and equitable infrastructure provisioning across school types and regions as a first-order lever schools control directly. **Policy:** results favor responsible-use curricula over access-restriction mandates, support dedicated AI-pedagogy modules in teacher-training accreditation, and highlight digital-equity provisioning as a distributive-justice concern with direct academic-outcome stakes. **Ethical considerations:** four concerns are directly relevant — potential bias/fairness gaps for multilingual populations; persistent hallucination risk given secondary students' relatively undeveloped independent-verification skills (motivating Ethical AI Awareness as a moderator rather than decorative construct); privacy given students' minor status, requiring parental/guardian consent in any real deployment; and the need for continued human (teacher) oversight as a complement to, not replacement for, awareness-building.

LIMITATIONS AND FUTURE RESEARCH

The primary limitation, stated without qualification, is that Results were computed on a simulation-calibrated rather than real dataset; every statistic is a genuine, reproducible computation, but on synthetic data whose structural relationships were specified in advance, so hypothesis support should be read as "supported under specification," not "empirically confirmed." Standard limitations that would also apply to a field-validated version include the cross-sectional design's inability to establish causal direction, geographic/institutional boundedness of the proposed sampling frame, self-report measurement susceptibility to social-desirability bias, and the rapidly evolving nature of the underlying technology.

Priority future-research directions: (1) conduct a genuine IRB-approved field administration of the instrument, replacing simulated with real data — the single highest-priority next step; (2) triangulate self-report Academic Performance with objective measures (grades, test scores); (3) adopt a longitudinal design to test engagement–antecedent causal direction; (4) add an experimental component (e.g., ethical-AI-awareness training vs. matched control) to support causal inference on the moderation finding; (5) test measurement invariance, not just structural-path equivalence, across student and teacher subsamples; (6) extend multi-group analysis to gender, school type, and region as additional moderators.

CONCLUSION

This study proposed and estimated the LLM-Enhanced Secondary Education Framework (LSEF), linking five antecedents to four learning outcomes through student engagement, with ethical AI awareness moderating the engagement–performance relationship. Using a fully disclosed, simulation-calibrated dataset, the LSEF measurement model achieved strong reliability and validity, and its structural specification produced coherent, well-estimated relationships consistent with the reviewed literature. The contribution of this work is not a claim of confirmed empirical fact about real secondary classrooms — that would require the field validation outlined above — but a theoretically integrated, psychometrically sound, fully reproducible framework and instrument that closes five specific literature gaps and is demonstrably ready for that field validation. As LLMs continue to diffuse into secondary classrooms faster than policy or evidence can keep pace, frameworks of this kind, once field-tested, offer a concrete route to evidence-based rather than reactive integration of generative AI into secondary schooling.

DATA AVAILABILITY, FUNDING, AND CONFLICT OF INTEREST

The simulation-calibrated dataset, data-generating script, pre-specified true parameters, and all analysis/figure code are available from the corresponding author on reasonable request, to support a subsequent field-validation study. This research received no external funding. The author declares no conflict of interest.

References

1. Chin, W. W., Marcolin, B. L., & Newsted, P. R. (2003). A partial least squares latent variable modeling approach for measuring interaction effects. *Information Systems Research*, 14*(2), 189–217.
2. Dindar, M., Suorsa, A., Hermes, J., Karppinen, P., & Näykki, P. (2021). Comparing technology acceptance of K-12 teachers with and without prior experience of learning management systems. *Journal of Computer Assisted Learning*.
3. Frontiers in Education. (2026). A systematic critical review of generative AI's impact on authorship, pedagogy, and integrity (2023–2025). *Frontiers in Education*.
4. Henseler, J., & Chin, W. W. (2010). A comparison of approaches for the analysis of interaction effects between latent variables using PLS path modeling. *Structural Equation Modeling*, 17*(1), 82–109.
5. Hunt, D., et al. (2026). Evaluating the performance of 3 large language models in higher-education essay-like assessments. *American Journal of STEM Education*.

6. MDPI Electronics. (2025). Exploring the educational applications of large language models: A systematic review and topic analysis. *Electronics, 14*(23), 4683.
 7. MDPI Information. (2025). Generative AI and academic integrity in higher education: A systematic review and research agenda. *Information, 16*(4), 296.
 8. ScienceDirect / Computers and Education: AI. (2025). Large language models in education: A systematic review of empirical applications, benefits, and challenges.
 9. ScienceDirect / Computers and Education: AI. (2024). Examining teachers' behavioural intention of using generative AI tools based on the extended technology acceptance model.
 10. Springer / Education and Information Technologies. (2025). Assessing ChatGPT acceptance and use in education: A comparative study among German-speaking students and teachers.
 11. Strzelecki, A. (2023). Students' acceptance of ChatGPT for education: An extended UTAUT2 model. *Innovative Higher Education*.
 12. Xu, J., & Deng, Q. (2024). Exploring Chinese secondary school students' acceptance of live video-streamed teaching platforms in EFL class. *Behavioral Sciences, 14*(7), 593.
 13. *Note: this is a curated, real, verifiable reference subset; expanding to full Q1-submission depth would require additional structured database searching beyond this draft's scope.*.
- .