2026

EJEL Volume 24, Issue 2

Editors

Heinrich Söbke

Published by Academic Publishing International Limited
Curtis Farm, Kidmore End, Nr Reading, RG4 9AY, United Kingdom
karen.harris@academic-publishing.org

eISSN: 1479-4403

EJEL Volume 24, Issue 2

Machine Learning in Art Teacher Education: A Comparative Analysis and Student Perceptions

Botagoz Kystaubayeva[1], Gulmira Mailybaeva[1], Kairat Dzhanabaev[2], Ainur Ansabayeva[1], Elmira Kydyrbekova[3] and Aivar Sakhipov[4]

[1]Department of Teaching and Upbringing Methods, Zhetysu university named after I. Zhansugurov, Taldykorgan, Kazakhstan

[2]Department of Physical and Basic Military Training, Zhetysu university named after I. Zhansugurov, Taldykorgan, Kazakhstan

[3]School of Arts and Humanities, Astana International University, Kazakhstan

[4]School of Software Engineering, Astana IT University, Kazakhstan

b.kystaubayeva@zu.edu.kz
g.mailybayeva@zu.edu.kz
dzanabaevkajrat072@gmail.com
ainuransabayeva@gmail.com (corresponding author)
e20250707@gmail.com
aivar.sakhipov@astanait.edu.kz

https://doi.org/10.34190/ejel.24.2.4313

Abstract: Amid the global push for digital transformation in higher education, there is a critical need for objective, scalable assessment tools in subjective disciplines like visual arts. Modern teacher education increasingly integrates intelligent technologies, yet the application of machine learning (ML) for formative assessment in art education remains underexplored. While ML offers scalable feedback, its capacity to evaluate subjective creativity remains contested. The study aims to examine the technical accuracy of a CNN-based model trained on a local dataset of 300 archived projects, compared to instructor evaluations, and to analyze how future teachers (N = 180) perceive algorithmic feedback in assessment contexts. A mixed-methods design was employed using a highly reliable survey instrument (Cronbach's α = .925) and comparative scoring analysis across four key dimensions: Technique, Composition, Color, and Creativity. Results indicate that the model aligns strongly with human assessments on technical execution (r = .426, p < .001), and moderate alignment for Composition (r = .430, p < .001) and weaker alignment for Color (r = .327, p < .001), while correlations for Creativity were notably weaker (r = .181, p = .015), indicating persistent limitations in modeling abstract artistic intent. ANOVA results revealed that students' digital literacy significantly predicts their trust in the system (F = 3.547, p = .031) and willingness to use it (F = 8.476, p < .001). Furthermore, discrepancy analysis indicated systematic divergence across proficiency levels, with the model exhibiting increasing underestimation for highly proficient students, particularly in cases involving stylistic deviation or non-standard cultural expression. The findings suggest that while the algorithm provides consistent, transparent scoring that enhances assessment literacy, it lacks the sensitivity to evaluate high-level originality due to standardization bias. This study contributes to the field by empirically demonstrating the "accuracy–creativity trade-off" in ML-based art assessment and by validating a hybrid assessment framework that balances algorithmic precision with pedagogical intuition. The study concludes that ML tools should function as "human-in-the-loop" support systems rather than autonomous graders, fostering critical reflection and digital competence in future educators.

Keywords: Pre-service primary teachers, Visual arts education, Teaching practices, Artificial intelligence, Machine learning

1. Introduction

Modern teacher education is undergoing rapid digital transformation, requiring a reconsideration of both the content of teacher preparation and the methods used to assess students' learning outcomes (Cai, 2025; Bedir and Freedman, 2024). As digital technologies become integral to contemporary pedagogical practice, intelligent systems are increasingly explored as tools for supporting assessment and formative feedback (Buckingham Shum et al., 2023). However, while machine learning (ML) approaches are widely applied in structured academic domains, their use in evaluating students' visual and creative work remains limited (Patterson et al., 2024). This gap is particularly pronounced in pre-service teacher education, where creative assignments are central to professional formation, yet assessment practices continue to rely heavily on subjective instructor judgment (Buckingham Shum et al., 2023; Patterson et al., 2024).

The assessment of visual and creative work is challenging due to the difficulty of formalizing evaluative criteria such as originality, composition, color harmony, and expressive intent. This subjectivity complicates consistency

Cite this article: Kystaubayeva, B. et al. 2026. "Machine Learning in Art Teacher Education: A Comparative Analysis and Student Perceptions", *Electronic Journal of e-Learning*, 24(2), pp 01-16, https://doi.org/10.34190/ejel.24.2.4313

and transparency in evaluation, constrains opportunities for structured student self-reflection, and increases instructors' assessment workload. Recent advances in computer vision and machine learning suggest that algorithmically supported approaches may contribute to formative assessment by providing consistent, rubric-aligned feedback on selected visual features. Within a human-in-the-loop framework, such outputs are designed to support professional pedagogical judgment rather than to replace it (Zheng et al., 2025; Patterson et al., 2024).

In this study, algorithmic assessment refers to the generation of criterion-based formative feedback through machine learning models that analyze predefined visual and semantic features of student work. It does not involve automated summative grading or the delegation of evaluative authority to algorithms. Engagement with such feedback is closely related to digital pedagogical competence, understood here as the ability of future teachers to critically engage with digital tools, reflect on their pedagogical affordances and limitations, and apply them responsibly within educational practice.

The present study explores the potential of algorithmically supported formative assessment in art-oriented teacher education. Specifically, it investigates the use of a machine learning model for evaluating digital art projects produced by pre-service primary school teachers. The scope of the study is deliberately limited to undergraduate teacher education and to the analysis of visual characteristics of digital artworks alongside keyword-based thematic features derived from project descriptions. Other modalities, broader learning analytics, and automated summative grading fall outside the focus of the research.

The study was conducted within the bachelor-level academic program 6B01301 "Pedagogy and Methods of Primary Education," implemented at Zhetysu University named after I. Zhansugurov and Abai Kazakh National Pedagogical University in Kazakhstan. The program is delivered within faculties responsible for teacher education and focuses on pedagogical reflection, diagnostic competence, and the integration of digital educational technologies. The analyzed data consist of final digital art projects completed within a course designed to foster students' creative, digital, and pedagogical competencies. Before outlining the theoretical framework and methodology, the main contributions of this study are summarized as follows:

- The study demonstrates how algorithmically supported formative assessment can be applied to creative tasks in teacher education, clearly distinguishing it from automated summative grading and positioning it as a human-in-the-loop support tool.
- It proposes a hybrid ML-based assessment model that combines convolutional neural networks, keyword-based thematic clustering, and rubric-aligned feedback for evaluating digital art projects.
- It provides empirical evidence from an authentic teacher education context by comparing ML-generated feedback with instructor evaluations and analyzing pre-service teachers' perceptions of and trust in automated feedback.
- It examines how engagement with algorithmic feedback supports assessment literacy, critical reflection, and understanding of algorithmic decision-making, without claiming evidence of competence development.

The study involved volunteer participants from two Kazakhstani universities and was fully integrated into the regular educational process. It did not interfere with students' academic trajectories or grading decisions, involved no sensitive or identifiable data, and was conducted under pedagogically neutral conditions.

Guided by these aims, the following research questions (RQ) were formulated:

RQ1: To what extent do the assessments generated by the machine learning model correlate with instructor evaluations of students' digital projects based on visual criteria?

RQ2: How do prospective primary school teachers perceive automated formative evaluation of visual work, and to what degree are they willing to trust such forms of feedback?

RQ3: To what extent does algorithmic assessment contribute to the development of elements of digital pedagogical competence, such as critical reflection, understanding of algorithmic principles, and self-assessment skills?

2. Literature Review

2.1 Machine Learning in Educational Assessment: Opportunities and Epistemic Limits

Machine learning is increasingly employed in educational assessment to provide scalable and consistent feedback, particularly in domains where evaluation criteria can be precisely formalized (Samuel, 2024; U.S. Department of Education, 2023). In such contexts, ML-based systems show substantial alignment with expert judgment, especially when trained on well-annotated datasets and applied to rule-governed outputs (Kusuma et al., 2022; Misgna et al., 2024).

This alignment weakens when assessment relies on pedagogical interpretation rather than formal pattern recognition. ML models optimize statistical regularities, whereas educational judgment in teacher education depends on context, intent, and professional expertise (Hopfenbeck et al., 2023; Samuel, 2024). Accordingly, automated scoring is increasingly positioned as rubric-aligned, human-in-the-loop support rather than a replacement for instructor judgment (Xu et al., 2025).

Explainability improves transparency and learner trust, as interpretable rationales are perceived more positively than opaque scores (Conijn, Kahr and Snijders, 2023; Chai et al., 2024). However, explainability does not grant access to pedagogical intent, meaning-making, or creative reasoning, even when instance-level explanations are provided (Rachha and Seyam, 2023; Saqr and López-Pernas, 2024; Khosravi et al., 2022). Accordingly, recent literature emphasizes mitigation rather than elimination of epistemic limits, through rubric alignment, provisional feedback framing, and explicit preservation of instructor authority (Xu et al., 2025; Conijn, Kahr and Snijders, 2023).

Responsible deployment therefore requires pedagogically grounded and equity-oriented frameworks, as automated assessment can otherwise reinforce bias and narrow evaluative norms (Dringó-Horváth, Rajki and Nagy, 2025; Miao and Cukurova, 2024). Interdisciplinary collaboration between educators, assessment specialists, and data scientists is widely identified as essential for ethical and educational validity (Guo et al., 2024). In teacher education, ML-based assessment thus plays a dual role: supporting learning while shaping future teachers' understanding of assessment practices and their limits.

2.2 Automated Assessment in Visual Arts: Technical and Conceptual Limits

Despite advances in educational ML, applications in visual and creative domains remain structurally constrained. Visual artworks frequently encode meaning through symbolism, abstraction, emotional expression, and stylistic deviation, dimensions that resist formal standardization. Convolutional neural networks show moderate to high agreement with expert ratings when evaluation targets formal visual properties such as composition, symmetry, or color distribution (Cropley and Marrone, 2025; Patterson et al., 2024). However, performance declines systematically when evaluation depends on originality, abstraction, or expressive intent (Cropley and Marrone, 2025; Patterson et al., 2024). Recent studies attempt to approximate creativity and emotional expression through indirect proxies, including novelty detection, stylistic divergence from training distributions, or multimodal alignment between images and short textual descriptors (Messer, 2024; Spee et al., 2023). These approaches remain fundamentally indirect and do not constitute pedagogical interpretation, as they lack access to artistic intent, emotional meaning, and culturally situated symbolism (Mazzone and Elgammal, 2019).

Empirical studies show systematic underestimation of works that deviate from dominant visual regularities, particularly those incorporating cultural symbolism or non-dominant aesthetic traditions (Spee et al., 2023; Zhang et al., 2025). Such divergence reflects a semantic gap rather than random noise. Models trained on visual regularities tend to interpret deviation as error because intent, context, and symbolic reference are not represented in visual feature space (Cetinic and She, 2022; Cibotaru, 2025). This limitation is architectural rather than empirical: expanding datasets may reduce variance but cannot resolve the absence of semantic understanding.

Cultural asymmetries further complicate automated assessment. Much of the literature is grounded in Western art education traditions that privilege individualist and modernist aesthetic norms (Crawford and Paglen, 2021). In contrast, Eastern and Central Asian artistic practices often emphasize culturally embedded symbolism, narrative continuity, and collective meaning-making (Bao et al., 2016). When models trained predominantly on Western datasets are applied in such contexts, stylistic deviation may be misclassified as low quality rather than culturally situated expression, a concern particularly relevant to the present study conducted in Kazakhstan (Coeckelbergh, 2023).

Pedagogical usefulness is further constrained by opacity. When automated systems output scores without interpretable justification, learners struggle to relate feedback to artistic intent, undermining trust and formative value (Nazaretsky et al., 2025). Although multimodal approaches combining visual and textual inputs show promise, most remain experimental and are rarely embedded in authentic instructional settings.

Overall, ML systems can reliably quantify visual structure but remain poorly suited to evaluating expressive deviation. Effective use in arts education therefore requires explicit role delimitation and instructor mediation, positioning algorithmic output as analytic support rather than autonomous judgment, particularly where meaning, symbolism, and originality are central (Fong and Schallert, 2023; Grájeda et al., 2024).

Student perceptions play a critical role in the educational legitimacy of AI-based assessment systems. Across higher education contexts, students tend to adopt a pragmatic but cautious stance toward automated evaluation: structured feedback is valued for clarity and consistency, while interpretive authority is consistently reserved for human instructors, especially in creative domains (Holmes, Bialik and Fadel, 2019; Chan and Hu, 2023; Tierney, Peasey and Gould, 2025). Trust and acceptance are closely linked to students' digital and AI-related competence, with higher literacy associated with greater confidence in algorithmic feedback and its responsible use (Jin et al., 2025; Anand and Hu, 2024; Ng et al., 2023; Tenberga and Daniela, 2024). These findings suggest that learner engagement with AI assessment depends less on technical accuracy alone than on transparency, contextual framing, and pedagogical mediation.

2.3 Research Gaps and Motivation for the Present Study

Despite growing interest in ML-assisted assessment, several gaps remain. First, most studies on creative ML assessment are conducted in laboratory or experimental settings, limiting pedagogical relevance and transferability to authentic coursework (Bulut et al., 2024; Gunasekara and Saarela, 2025; Küchemann et al., 2025). Second, learner-facing explainability remains underdeveloped, restricting opportunities for reflective engagement with assessment logic (Khosravi et al., 2022; Gunasekara and Saarela, 2025). Third, empirical validation within teacher education is scarce, with few studies combining instructor–model performance comparison, diagnostic failure analysis, and student perception data within the same instructional context (Jankowsky and Schroeders, 2022). Finally, ethical and cultural sensitivity issues remain insufficiently addressed, particularly for non-Western and stylistically diverse student populations (Chinta et al., 2024; Ferrara, 2024; Fu and Weng, 2024).

These gaps motivate the present study, which examines a rubric-aligned ML assessment pipeline for digital art projects within pre-service teacher education. By combining quantitative performance analysis, structured interpretation of failure modes, and student perception data in an authentic classroom context, the study aims to clarify both the capabilities and the structural limits of algorithmic assessment in creative domains.

By integrating technical performance analysis with epistemological and pedagogical perspectives, this review contributes to the emerging theoretical understanding of why ML-based assessment remains structurally constrained in visual arts education.

3. Materials and Methods

3.1 Research Design

The study employed a mixed-methods design combining quantitative and qualitative components. This approach enabled analysis of both the correlation between instructor-assigned grades and machine-generated scores, and students' perceptions of automated assessment within teacher education. Quantitative analysis was based on a four-criterion rubric, while qualitative data were collected through an anonymous online survey. This combination provided a comprehensive view of the model's performance and its educational implications.

The research was integrated into a regular academic course focused on developing digital visual competence among prospective primary school teachers. Student-created digital artworks, submitted as final assignments, served as the dataset. All instructor evaluations were completed before applying the machine learning model, ensuring that academic outcomes were not influenced and that the results reflect independent analysis. Quantitative data were analyzed using Pearson correlation coefficients, mean absolute error (MAE), and one-way ANOVA, while qualitative responses were examined through descriptive content analysis.

3.2 Participants and Educational Context

The study involved 180 undergraduate students enrolled in the academic program 6B01301 "Pedagogy and Methods of Primary Education" at Zhetysu University named after I. Zhansugurov and Abai Kazakh National Pedagogical University. All participants were taking a course in digital art, during which they completed their individual final projects using raster and vector graphic editors. These 180 student works formed the core dataset for model testing.

Student artworks were collected after course completion via the institutional learning management system (LMS), following the official recording of all course grades. Participants were undergraduate pre-service teachers in their second to fourth year of study, and the analyzed projects corresponded to the final assessment of the digital art course within the academic semester.

To train the machine learning model, an additional dataset of 300 archived projects from previous cohorts was used. These earlier works had been evaluated using the same four-criterion rubric by four instructors specializing in digital art education. To assess the consistency of human scoring, inter-rater reliability was calculated on a random subset of 60 projects evaluated independently by pairs of instructors. The average Cohen's kappa coefficient across all criteria was 0.72, indicating substantial agreement. All evaluations followed a collaboratively developed rubric. Prior to assessment, calibration sessions were held to ensure consistent interpretation of the criteria and to reduce subjectivity.

All new student projects were independently assessed both by instructors and by the machine learning model using identical evaluation criteria. After the evaluation phase, an anonymous online survey was administered to gather students' perceptions of the system's fairness, clarity, and relevance to their future teaching practice.

3.3 Evaluation Rubric and Analysis Structure

Assessment was conducted using a four-criterion rubric covering key aspects of the visual product: composition, color scheme, technical execution, and creativity. Each criterion was rated on a 5-point scale. The rubric was developed in consultation with instructors to ensure standardization and was applied in both human evaluation and machine processing. The structure of the rubric is presented in Table 1.

Table 1: Evaluation Rubric for Digital Art Projects

Criterion	Description
Composition	Visual balance and logical arrangement of elements in the image, including spatial organization and alignment
Color Scheme	Harmony, contrast, and emotional tone of selected colors, including consistency of color relationships and overall visual coherence
Technical Execution	Accuracy and clarity in the use of digital tools and techniques, including line precision, layering, and resolution
Creativity	Originality and divergence from typical patterns within the task constraints, reflected in non-standard visual solutions

The evaluation rubric was grounded in widely accepted formal criteria used in visual arts education, focusing on observable compositional, chromatic, and technical features (e.g., spatial balance, color contrast, and tool accuracy) rather than culturally specific symbolic interpretations. Creativity was operationalized as originality and divergence from typical patterns within the task constraints, avoiding references to stylistic canons or culturally bound aesthetic norms. This design allowed consistent application across student works reflecting diverse artistic traditions and cultural backgrounds.

The model produced outputs on a continuous numerical scale, which were rounded to one decimal place and then converted into rubric scores using predefined intervals. For instance, values from 1.0 to 1.4 corresponded to a score of 1, values from 1.5 to 2.4 to a score of 2, and so forth.

3.4 Model Architecture and Training

The machine learning model was developed using a supervised learning approach and trained on labeled data consisting of digital images and instructor-assigned scores. To provide limited contextual grounding, the model incorporated a keyword-based thematic grouping mechanism. Students selected keywords from a predefined list, allowing each submission to be associated with a thematic cluster of previously evaluated works. This approach offered minimal contextual reference while remaining aligned with rubric-based assessment constraints. The overall task was framed as a multivariate regression problem, aimed at predicting four continuous values corresponding to the rubric criteria.

A shallow convolutional neural network (CNN) architecture was intentionally selected due to the moderate size of the training dataset (300 archived projects). This design balanced feature expressiveness with reduced overfitting risk and improved transparency for educational use. Three convolutional blocks were sufficient to capture mid-level visual features relevant to rubric-based criteria such as composition, color distribution, and technical execution.

A convolutional neural network was implemented to extract visual features from student artworks and map them to rubric-based scores. The architecture comprised three convolutional blocks followed by fully connected layers and an output layer with four neurons corresponding to the rubric criteria. Input images were normalized and resized to 512×512 pixels.

Model training was conducted using the mean squared error (MSE) as the loss function and the Adam optimizer. Training used an 80/20 train–validation split with early stopping based on validation loss, a batch size of 32, and a maximum of 50 epochs. Standard data augmentation techniques were applied to improve generalization. A detailed breakdown of the architecture is provided in Table 2.

Table 2: CNN Model Architecture

Step	Layer Type	Parameters	Output Dimensions
1	Conv2D + ReLU	3×3 kernel, 32 filters	512×512×32
2	MaxPooling	2×2 window	256×256×32
3	Conv2D + ReLU	3×3 kernel, 64 filters	256×256×64
4	MaxPooling	2×2 window	128×128×64
5	Dense	128 neurons	–
6	ReLU	–	–
7	Output Layer	4 neurons	–

Figure 1 presents a schematic overview of the assessment pipeline, illustrating image and keyword input, feature extraction, inference, and rubric-aligned feedback generation, with optional visual interpretability overlays.

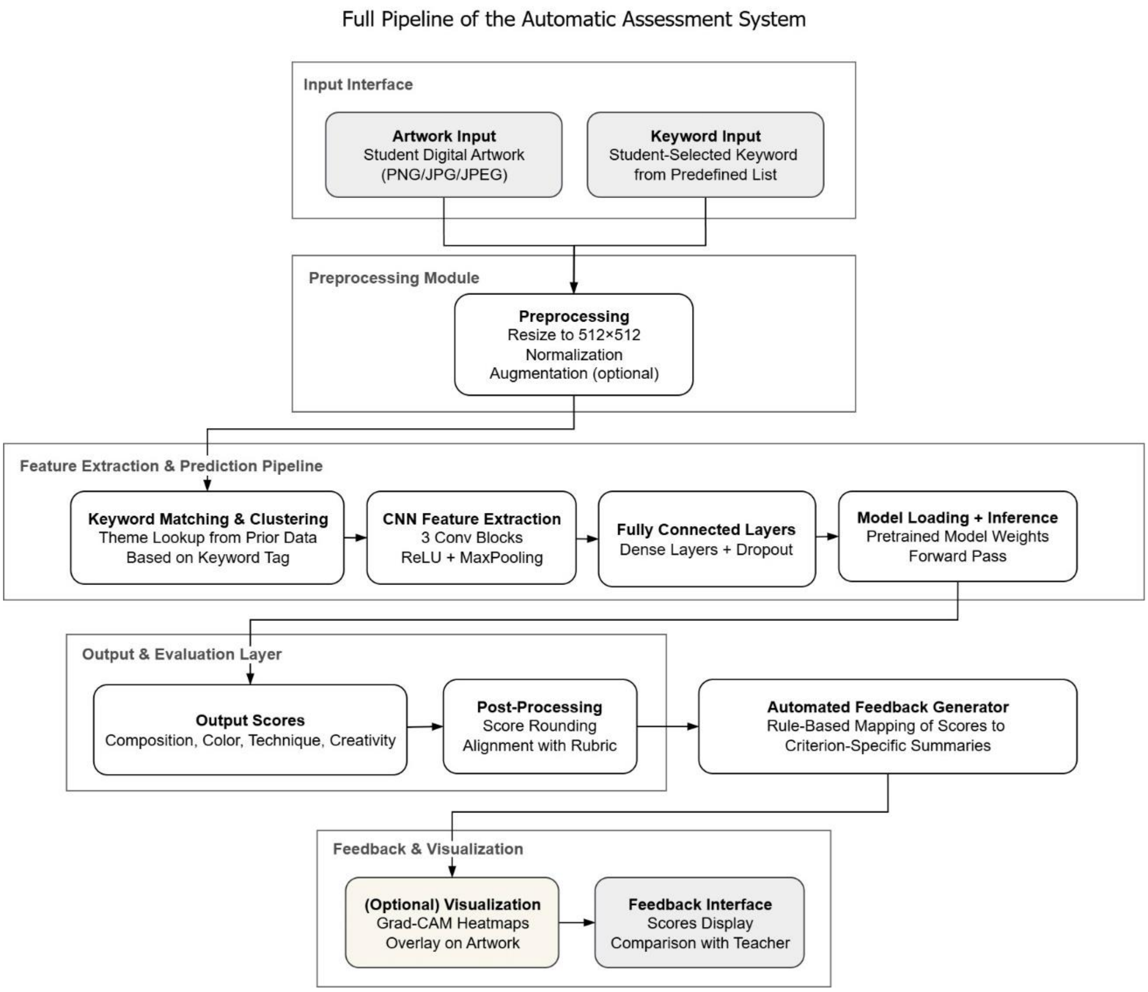


Figure 1: Automatic Assessment Pipeline for Student Digital Artwork

3.5 Comparative Analysis of Scores

Each of the 180 new student works was independently evaluated by an instructor and by the machine learning model. Discrepancies between the two scoring systems were analyzed using Pearson correlation coefficients, as well as metrics such as mean absolute error (MAE) and the proportion of matches within a one-point margin. Additional attention was given to cases where the difference between the two evaluations was two points or greater. These outliers were examined in detail to identify possible sources of divergence, such as the use of unconventional color palettes, abstract composition, expressive symbolism, or other artistic elements that may not have been well represented in the training data. The analysis revealed key strengths of the model, including high agreement on technical criteria, alongside notable limitations, particularly reduced sensitivity to conceptual originality.

3.6 Survey Instrument Reliability and Analysis of Student Perceptions

Following the completion of the evaluation process, students participated in an anonymous online survey that included both Likert-scale and open-ended questions. The survey focused on levels of trust in the results, perceptions of objectivity, impact on motivation, and willingness to use similar tools in future teaching practice.

Prior to inferential analysis, the internal consistency of the survey instrument was evaluated. The Digital Literacy Scale (Q1–Q3) was developed specifically for this study context to measure domain-relevant self-efficacy rather than general digital skills. It aggregates perceived competence in three key areas: technical proficiency with editing tools, theoretical understanding of algorithms, and prior practical experience with AI. The scale demonstrated excellent reliability (Cronbach's $\alpha = .912$). Participants were stratified into three levels based on a summative index (Range: 3–15) established according to Likert scale anchor points: Low Literacy ($n = 49$, scores 3–8) reflected average responses below the neutral midpoint; Medium Literacy ($n = 61$, scores 9–11) represented functional proficiency centering around the neutral value; and High Literacy (n = 70, scores 12–15) corresponded to consistently high confidence ratings averaging 4 or above.

The Student Perceptions Scale (Q4–Q11), capturing trust, perceived objectivity, clarity, emotional response, and willingness to use AI-based assessment tools, also showed high internal consistency ($\alpha = .925$; $\omega = .927$), confirming the suitability of the instrument for group comparisons and subsequent inferential analyses.

Responses to the open-ended questions were subjected to content analysis, with coding carried out independently by two researchers. The resulting thematic structure was organized into four categories: perceived fairness, emotional response, pedagogical applicability, and critical attitudes toward algorithmic assessment.

3.7 Ethical Considerations

This study was part of everyday teaching practice at the university and followed basic institutional ethical principles normally used in low-risk educational studies. Teaching and grading were conducted in the usual way for the course, with instructors responsible for evaluating student work. Only after the course had ended were selected works used to examine the performance of the machine learning model in a post-hoc and purely experimental manner, aimed at analysis and formative insight rather than evaluation. The model was trained and tested using only anonymized student work, with all identifying details removed. Participation in the survey was voluntary and anonymous. As the study had no effect on teaching or student outcomes, formal ethical approval was not required.

4. Results and Findings

4.1 Comparative Analysis of Instructor and Model Scores

Agreement between instructor and model scores was examined for 180 student digital artworks. The evaluation focused on four rubric criteria: composition, color scheme, technical execution, and creativity. The results are presented in Table 3. Technical Execution showed the strongest alignment between instructor and model scores ($r = .426$, $p < .001$), with the lowest error values (MAE = 0.48; MSE = 0.85) and comparable mean scores (Instructor M = 4.14; Model M = 4.01), indicating stable performance on formally defined visual features.

Table 3: Agreement Between Instructor and Model Scores by Criterion

Criterion	Instructor Mean (SD)	Model Mean (SD)	MAE	MSE	Pearson r
Technical Execution	4.14 (0.80)	4.01 (0.90)	0.48	0.85	.426***
Creativity	3.98 (0.97)	3.67 (1.17)	1.03	1.97	.181*
Composition	4.03 (0.91)	3.79 (1.02)	0.62	1.12	.430***
Color Scheme	3.90 (0.97)	3.72 (1.06)	0.74	1.41	.327***

Note. * $p < .05$, *** $p < .001$

Creativity exhibited the largest divergence, with the highest absolute error (MAE = 1.03; MSE = 1.97) and a moderate correlation ($r = .181$, $p = .015$). Model scores were consistently lower than instructor ratings (Model M = 3.67 vs. Instructor M = 3.98), indicating systematic underestimation rather than random disagreement. Composition ($r = .430$; MAE = 0.62) and Color Scheme ($r = .327$; MAE = 0.74) showed intermediate agreement, with the model assigning consistently lower mean scores than instructors, suggesting conservative estimation for partially interpretive visual features.

Analysis by digital literacy level revealed systematic divergence patterns. While the model followed overall performance trends, it showed a ceiling effect for highly proficient students. In the High literacy group, the instructor mean for Creativity (M = 4.27) exceeded the model mean (M = 3.89) by over one point, whereas the gap was smaller in the Low literacy group (Instructor M = 3.16; Model M = 3.43). This indicates that the model tends to underestimate highly original work, particularly among highly proficient students, interpreting stylistic deviation as error rather than creative intent.

Overall, model performance differed systematically across rubric dimensions. Criteria grounded in formally defined visual properties showed stronger alignment, whereas creativity and expressive intent exhibited higher absolute error and consistent downward bias. This pattern, reported in prior work on visual arts and automated scoring (Cetinic and She, 2022; Cropley and Marrone, 2025; Misgna et al., 2024; Patterson et al., 2024), indicates that correlation alone may mask systematic score compression. Taken together, the results delineate a functional boundary: the model supports structured assessment but remains limited in evaluating originality.

4.2 Score Distributions and Divergence Analysis

To further examine agreement patterns, the distributions of instructor and model scores were visualized separately for each rubric criterion. Figure 2 presents overlaid histograms comparing the score distributions of the instructor and the model across all four criteria.

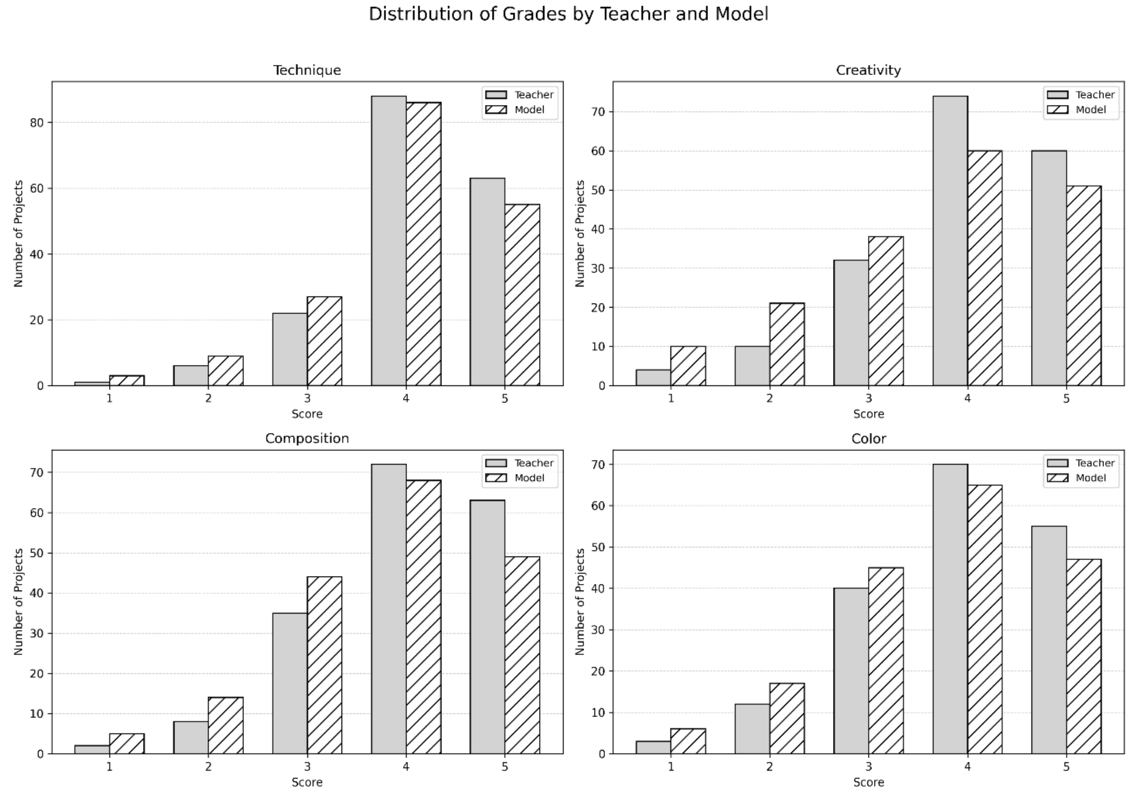


Figure 2: Comparative Post-Rounded Distribution of Instructor and Model Scores by Criterion

Across all criteria, model score distributions were shifted downward relative to instructor ratings, most prominently for Creativity. Smaller but consistent shifts were also present for Composition and Color Scheme, whereas distributions for Technical Execution showed closer alignment.

Analysis of cases with large discrepancies (≥ 2 points) revealed recurring patterns rather than isolated errors. According to the data in Figure 3, the most frequent sources of divergence were abstract composition (38%), unconventional color use (27%), symbolic or metaphorical content (22%), and mixed stylistic approaches (13%). These categories represent a stable set of divergence types across the dataset.

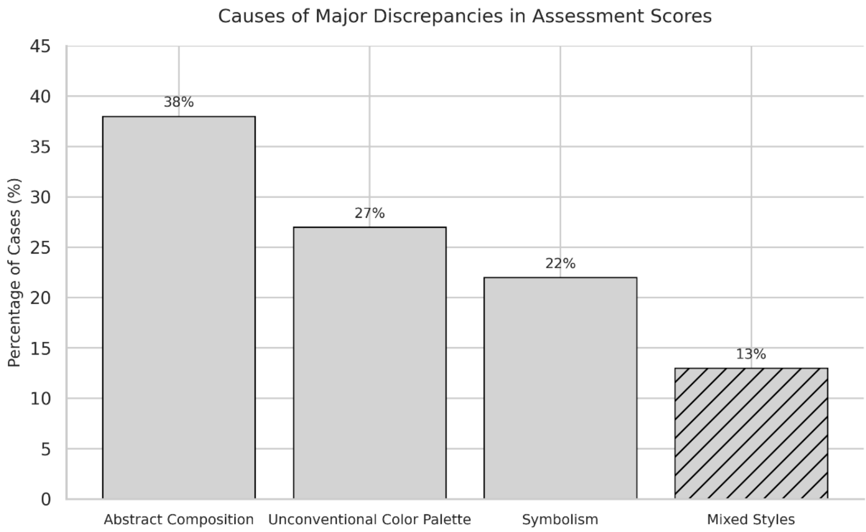


Figure 3: Sources of Major Score Discrepancies

Distributional patterns confirm that divergence follows recurring visual characteristics rather than random error, with the greatest dispersion observed for originality and non-standard visual expression.

4.3 Qualitative Evaluation and Sample-Wide Agreement

Qualitative inspection of selected cases revealed recurring patterns of disagreement. Divergences were most frequent in works employing abstract composition, expressive color contrasts, or minimalistic strategies, where instructors interpreted deviation as intentional, while the model treated it as deficiency.

Figure 4 presents average score deviations between model predictions and instructor evaluations for the full dataset (N = 180). The largest negative deviation is observed for the Creativity criterion (mean post-rounded deviation = –0.31), indicating a systematic tendency of the model to assign lower creativity scores relative to instructor judgments. Smaller but consistent negative deviations are also evident for Composition, Color, and Technique, suggesting a generally conservative scoring pattern across evaluation dimensions.

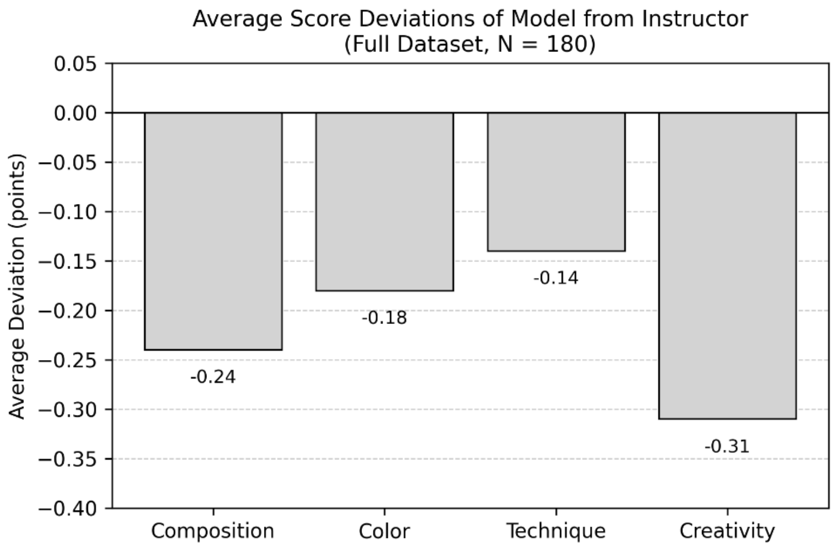


Figure 4: Average Post-Rounded Deviations Between Model and Instructor Evaluations

Combined qualitative and aggregate analyses show that agreement is highest for technically explicit criteria, while divergence increases for expressive and symbolic dimensions, supporting the use of the system as analytic support rather than autonomous evaluation.

4.4 Student Perceptions of the Technology

One component of the study focused on how students perceived the use of automated assessment during the learning process. An anonymous online survey was completed by 180 participants. The survey included Likert-scale items measuring trust in the model, perceived clarity and fairness of feedback, emotional response, and willingness to use similar systems in future teaching practice. Table 4 summarizes the distribution of respondents who agreed or strongly agreed with each statement.

Table 4: Distribution of Student Responses to Scaled Statements

Statement	Agree or strongly agree
The model evaluates student work more objectively than a human assessor.	114 (63.3%)
The model's scores are clear and understandable.	129 (71.7%)
I would consider using such a system in my future teaching practice.	105 (58.3%)
The absence of emotional nuance reduces the quality of feedback.	120 (66.7%)
The model helps clarify assessment criteria.	144 (80.0%)

Survey responses indicate that students viewed the system as transparent and helpful for understanding assessment criteria, while clearly recognizing its limitations in conveying emotional nuance. Open-ended responses (n = 48) emphasized improved clarity of criteria, conditional trust, limited emotional sensitivity, and cautious classroom applicability.

To examine whether digital literacy influenced student perceptions, respondents were stratified into Low (n=49), Medium (n = 61), and High (n = 70) levels based on their self-reported confidence in basic image editing and their understanding of algorithmic principles. Table 5 presents the descriptive statistics (Mean and Standard Deviation on a 5-point scale) and the results of a one-way ANOVA for each perception variable. The analysis confirmed statistically significant differences across all three constructs. Students with higher digital literacy reported significantly greater trust in the model, a stronger willingness to use the tool, and a clearer understanding of its logic. Post-hoc comparisons indicated that the "High" literacy group consistently rated the system more favorably than the "Low" literacy group, suggesting that technical competence is a key predictor of AI acceptance.

Table 5: Influence of Digital Literacy on Student Perceptions (Mean Scores and ANOVA Results)

Perception Variable	Mean (SD)			ANOVA Results
	Low Literacy (n = 49)	Medium Literacy (n = 61)	High Literacy (n = 70)	
Trust in the Model	3.37 (1.38)	3.69 (1.37)	4.00 (1.12)	$F(2, 177) = 3.55$, $p = .031$
Willingness to Use	2.76 (1.56)	3.54 (1.41)	3.81 (1.28)	$F(2, 177) = 8.48$, $p < .001$
Understanding of Logic	3.16 (1.38)	3.77 (1.35)	4.27 (0.96)	$F(2, 177) = 11.92$, $p < .001$

Note. Scores are based on a 5-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree). SD = Standard Deviation. Differences between groups are statistically significant at $p < .05$.

These results indicate a monotonic pattern: as digital literacy increases, students report higher confidence in the fairness and usefulness of algorithmic assessment, as well as greater comprehension of how model outputs are generated. At the same time, even among highly literate students, reservations regarding emotional nuance and creative interpretation remain present, underscoring the perceived need for continued instructor involvement. Overall, students perceived the system as a supportive assessment aid rather than a replacement for teacher judgment, valuing its structure while assigning responsibility for creative interpretation to instructors.

5. Discussion

5.1 Differential Model Performance Across Rubric Dimensions

The results demonstrate systematic differences in model performance across rubric dimensions. Alignment with instructor ratings is strongest for technically defined criteria such as Technical Execution and remains relatively high for Composition, where evaluation relies on formally operationalized visual features. Creativity follows a distinct pattern. Despite low correlations, absolute error is substantially higher and model scores are consistently lower than instructor ratings, indicating systematic underestimation rather than random disagreement. Reduced variance further suggests conservative model behavior when originality departs from familiar visual patterns. This boundary reflects the system's design: convolutional feature extraction and keyword aggregation support reliable assessment of technical and structural properties but provide limited access to expressive intent. Across metrics and samples, the findings support the use of the model as an analytic aid for structured criteria, with instructor judgment remaining central for creative evaluation.

5.2 Systematic Divergence, Failure Modes, and the Role of Instructor Judgment

Analysis of large score discrepancies shows that divergence between model and instructor evaluations follows stable failure modes linked to specific elements of the assessment pipeline rather than random error. Three mechanisms are consistently observed. First, a semantic gap emerges in keyword-based clustering: categorical prototypes represent an "average" visual solution, causing symbolic or metaphorical works to be penalized when their visual features diverge from expected patterns. Second, the shallow CNN architecture introduces a formalism bias. While effective for detecting mid-level features such as edge clarity and structural balance, it favors conventional compositional regularities and interprets abstract or non-standard structures as technical deficiencies rather than intentional rule-breaking. Third, regression-based scoring applies a distributional penalty to unconventional color schemes, treating palettes outside dominant training distributions as noise rather than expressive choice. Together, these mechanisms function as a normalization filter that compresses score variance and systematically pulls highly original work toward the mean.

These limitations define clear boundaries of autonomous algorithmic judgment. Agreement with instructor ratings is highest where assessment depends on visible and repeatable cues, such as technical execution, and declines when evaluation relies on sensitivity, emotional nuance, or symbolic meaning. Creativity therefore shows the largest absolute deviation and a consistent downward bias, even when correlations remain low. This pattern reflects the system's representational scope: the model processes formal visual features and recurring statistical patterns but has no access to intention, cultural reference, or affect. As a result, sensitivity, originality, and creativity are preserved through use design rather than algorithmic inference. Model outputs function as provisional indicators that structure attention around explicit criteria, while instructors retain interpretive authority by contextualizing feedback through analysis of intent, symbolism, and expressive trajectory. The system thus operates as analytic support for formal assessment, stabilizing technical evaluation while preserving human judgment over creative and contextual meaning

5.3 Pedagogical and Methodological Implications

The findings have several implications for teacher education practice. Algorithmic assessment can increase transparency and consistency in formative evaluation, particularly in domains characterized by subjective judgment. At the same time, such systems should function as complementary tools rather than substitutes for human evaluators (Li and Botelho, 2024). The results confirm that ML-based assessment is most effective for technical and structural criteria, while rubric-aligned feedback and keyword-based grouping can support awareness of creative and symbolic aspects without claiming interpretive authority.

Engaging with algorithmic feedback supports students' understanding of assessment criteria and facilitates reflective comparison between human and machine judgment, contributing to assessment literacy and digital awareness (Lawasi, Rohman and Shoreamanis, 2024). Rather than prescribing evaluation outcomes, the system encourages reflection on how formal criteria are operationalized and where algorithmic judgment reaches its limits. The study also identifies methodological directions for improvement, including expanded training datasets, clearer visualization of scoring logic, and the integration of explanatory comments, reinforcing the role of the system as a reflective aid rather than a prescriptive evaluator. Similar principles have been noted in related applications of ML for academic integrity and monitoring in digital learning environments (Sakhipov, Omirzak and Fedenko, 2025).

To support transparency, a dedicated interface was developed to visualize key stages of the assessment pipeline. The interface (Figure 5) presents score distributions, keyword tagging, and editable instructor inputs, illustrating how automated output and human judgment can be combined in a proof-of-concept configuration.

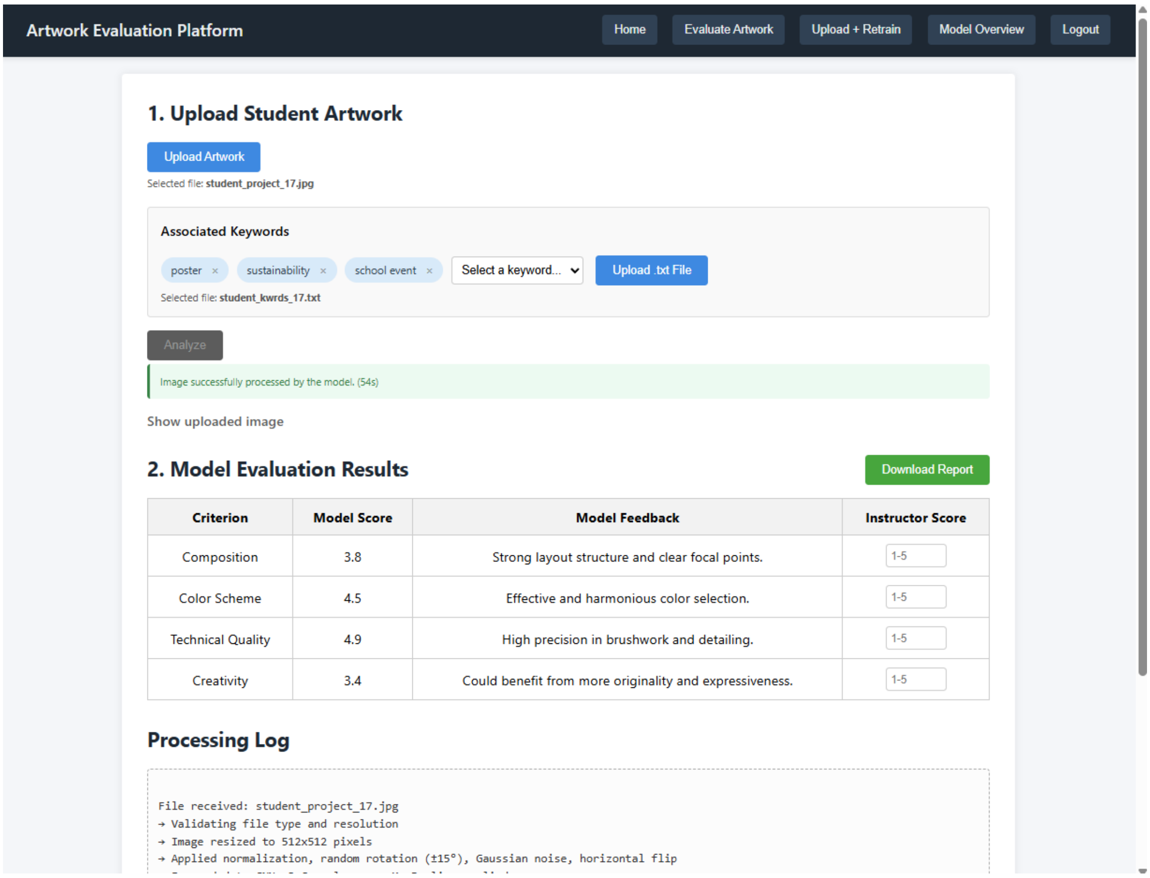


Figure 5: Interface of the Automated Formative Assessment System

5.4 Responses to the Research Questions

Regarding RQ1, the analysis showed relatively strong alignment between instructor evaluations and model scores for structurally defined criteria such as technical execution and composition, with low error rates and most scores falling within one point of instructor judgments. Performance declined for creative criteria involving originality and symbolic intent, indicating that while the model reliably assesses measurable visual features, interpretive evaluation remains dependent on instructor judgment.

For RQ2, survey results indicate a generally positive but cautious student stance toward automated feedback. Respondents valued clarity and consistency, while expressing concerns about limited emotional and contextual sensitivity. Trust and willingness to use the system varied significantly by digital literacy level, with higher literacy associated with greater confidence in the system's fairness and relevance. These differences are interpreted as associative rather than causal.

Regarding RQ3, the study does not provide evidence of learning gains or competence development. Instead, it documents increased awareness of assessment criteria and greater engagement in reflective comparison between human and algorithmic evaluations. In this study, algorithmic assessment thus contributes to reflective orientation and assessment literacy, rather than directly to pedagogical skill acquisition.

5.5 Study Limitations and Validity Considerations

This study, while informative, has several limitations. The dataset of 180 current and 300 archived projects supports comparative analysis but does not capture the full diversity of visual expression, particularly forms where meaning is conveyed through culturally specific, symbolic, or minimal visual strategies. The research was conducted within a single pre-service teacher education program in Kazakhstan and focused on one course type, which constrains generalization to other educational settings. In contrast to much automated assessment research that targets narrowly formalized criteria, and to creative ML studies typically conducted in laboratory or benchmark settings, this study examines model behavior in an authentic classroom context; this increases ecological validity but also introduces contextual variability that cannot be fully controlled. The model processes visual features and keyword-based thematic labels but lacks semantic and contextual understanding, resulting in reduced sensitivity to conceptually rich but visually sparse works; this limitation is architectural and cannot be resolved through dataset expansion alone. Survey findings rely on self-reported perceptions and are subject to response bias, while variables such as prior AI experience were not independently controlled. The cross-sectional design precludes claims about learning gains or competence development, and the study does not provide longitudinal or behavioral evidence of instructional impact.

5.6 Future Directions for Model Development and Research

The findings point to several directions for further work. While a multimodal framework that combines visual features with keyword grouping is already implemented, future research should examine richer forms of semantic integration and interpretability support. Expanding the training corpus may reduce variance in technical scoring but is unlikely to resolve limitations related to symbolic interpretation. Additional studies should test the model in other visually oriented disciplines and educational contexts, with adapted rubrics and comparative instructor baselines. Further research is also needed to evaluate how different feedback representations, including visual vignettes and process-level explanations, influence understanding of algorithmic judgment and support reflective engagement with assessment criteria.

6. Conclusion

This study examined a rubric-aligned machine learning pipeline for supporting the assessment of digital art projects in pre-service teacher education. Quantitative comparison with instructor scores addressed RQ1 by showing strong alignment for structured criteria such as technical execution and composition ($r \approx .43$; MAE ≈ 0.5), alongside systematic underestimation for creativity (MAE > 1.0), indicating a stable accuracy–creativity trade-off rather than random error. RQ2 was addressed through survey and inferential analysis, which showed generally positive but cautious perceptions of automated feedback, with trust and willingness to use the system varying significantly by digital literacy level and concerns directed primarily at creative scoring rather than technical criteria. RQ3 was addressed by documenting increased awareness of assessment criteria and reflective comparison between human and algorithmic judgments, while explicitly not providing evidence of competence development or learning gains. Across all research questions, the findings reinforce that the system functions as rubric-aligned triage and analytic support, augmenting but not replacing instructor judgment. Safe and responsible use depends on preserving instructor authority over interpretation, contextual meaning, and creative intent, while treating algorithmic output as provisional rather than evaluative. Future work should focus on multimodal extensions that incorporate semantic input, longitudinal designs that test instructional impact, and classroom-based deployments that examine how such systems can be integrated into formative feedback practices without constraining creativity.

AI Statement: During the preparation of this manuscript, the authors used Grammarly to enhance readability and language.

Ethics Statement: Ethical approval was not required, as the study involved no intervention, personal data, or impact on academic outcomes.

References

Anand, B. and Hu, Y. (2024). Play testing and reflective learning AI tool for creative media courses. In: *Proceedings of the 16th International Conference on Computer Supported Education – Volume 1: CSEDU*, pp.146–158. https://doi.org/10.5220/0012633600003693

Bao, Y., Yang, T., Lin, X., Fang, Y., Wang, Y., Pöppel, E. and Lei, Q. (2016). Aesthetic preferences for Eastern and Western traditional visual art: identity matters. *Frontiers in Psychology, 7*, 1596. https://doi.org/10.3389/fpsyg.2016.01596

Bedir, S. and Freedman, K. (2024). Integrating digital technologies and AI in art education: pedagogical competencies and the evolution of digital visual culture. *Participatory Educational Research, 11*, pp.57–79. https://doi.org/10.17275/per.24.94.11.6

Buckingham Shum, S., Lim, L.-A., Boud, D., Bearman, M. and Dawson, P. (2023). A comparative analysis of the skilled use of automated feedback tools through the lens of teacher feedback literacy. *International Journal of Educational Technology in Higher Education, 20(1)*, 40. https://doi.org/10.1186/s41239-023-00410-9

Bulut, O., Beiting-Parrish, M., Casabianca, J.M., Slater, S.C., Jiao, H., Song, D., Ormerod, C., Fabiyi, D.G., Ivan, R., Walsh, C., Rios, O., Wilson, J., Yildirim-Erbasli, S.N., Wongvorachan, T., Liu, J.X., Tan, B. and Morilova, P. (2024). The rise of artificial intelligence in educational measurement: opportunities and ethical challenges. *Chinese/English Journal of Educational Measurement and Evaluation*, *5*(3), 3. https://doi.org/10.59863/MIQL7785

Cai, Y. (2025). Integrating machine learning in art education: research framework and theoretical analysis approach. *International Journal of Advanced Research and Interdisciplinary Scientific Endeavours*, *2*, pp.537–542. https://doi.org/10.61359/11.2206-2515

Cetinic, E. and She, J. (2022). Understanding and creating art with AI: Review and outlook. *ACM Transactions on Multimedia Computing, Communications, and Applications*, *18*(2), 66. https://doi.org/10.1145/3475799

Chai, F., Ma, J., Wang, Y., Zhu, J. and Han, T. (2024). Grading by AI makes me feel fairer? How different evaluators affect college students' perception of fairness. *Frontiers in Psychology*, *15*, 1221177. https://doi.org/10.3389/fpsyg.2024.1221177

Chan, C.K.Y. and Hu, W. (2023). Students' voices on generative AI: perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, *20*, 43. https://doi.org/10.1186/s41239-023-00411-8

Chinta, S.V., Wang, Z., Yin, Z., Hoang, N., Gonzalez, M., Le Quy, T. and Zhang, W. (2024). *FairAIED: Navigating fairness, bias, and ethics in educational AI applications*. arXiv. Available at: https://arxiv.org/abs/2407.18745 [Accessed 10 December 2025].

Cibotaru, V. (2025). Is there computational creativity?. *AI & Society*. https://doi.org/10.1007/s00146-025-02708-w

Coeckelbergh, M. (2023). Narrative responsibility and artificial intelligence. *AI & Society*, *38*, pp.2437-2450. https://doi.org/10.1007/s00146-021-01375-x

Conijn, R., Kahr, P. and Snijders, C. (2023). The effects of explanations in automated essay scoring systems on student trust and motivation. *Journal of Learning Analytics*, *10*(1), pp.37–53. https://doi.org/10.18608/jla.2023.7801

Crawford, K. and Paglen, T. (2021). Excavating AI: the politics of images in machine learning training sets. *AI & Society*, *36*, pp.1105–1116. https://doi.org/10.1007/s00146-021-01162-8

Cropley, D.H. and Marrone, R.L. (2025). Automated scoring of figural creativity using a convolutional neural network. *Psychology of Aesthetics, Creativity, and the Arts*, *19*(1), pp.77–86. https://doi.org/10.1037/aca0000510

Dringó-Horváth, I., Rajki, Z. and Nagy, J.T. (2025). University teachers' digital competence and AI literacy: moderating role of gender, age, experience, and discipline. *Education Sciences*, *15*(7), 868. https://doi.org/10.3390/educsci15070868

Ferrara, E. (2024). Fairness and bias in artificial intelligence: a brief survey of sources, impacts, and mitigation strategies. *Sci*, *6*(1), 3. https://doi.org/10.3390/sci6010003

Fong, C. and Schallert, D. (2023). "Feedback to the future": advancing motivational and emotional perspectives in feedback research. *Educational Psychologist*, *58*(1), pp.1–16. https://doi.org/10.1080/00461520.2022.2134135

Fu, Y. and Weng, Z. (2024). Navigating the ethical terrain of AI in education: a systematic review on framing responsible human-centered AI practices. *Computers and Education: Artificial Intelligence*, *7*, 100306. https://doi.org/10.1016/j.caeai.2024.100306

Grájeda, A., Córdova, P., Córdova, J.P., Laguna-Tapia, A., Burgos, J., Rodríguez, L. and Sanjinés, A. (2024). Embracing artificial intelligence in the arts classroom: understanding student perceptions and emotional reactions to AI tools. *Cogent Education*, *11*(1). https://doi.org/10.1080/2331186X.2024.2378271

Gunasekara, S. and Saarela, M. (2025). Explainable AI in education: techniques and qualitative assessment. *Applied Sciences*, *15*(3), 1239. https://doi.org/10.3390/app15031239

Guo, S., Latif, E., Zhou, Y., Huang, X. and Zhai, X. (2024). *Using generative AI and multi-agents to provide automatic feedback*. arXiv. Available at: https://arxiv.org/abs/2411.07407 [Accessed 10 December 2025].

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

Hopfenbeck, T.N., Zhang, Z., Sun, S.Z., Robertson, P. and McGrane, J.A. (2023). Challenges and opportunities for classroom-based formative assessment and AI: a perspective article. *Frontiers in Education*, *8*, 1270700. https://doi.org/10.3389/feduc.2023.1270700

Jankowsky, K. and Schroeders, U. (2022). Validation and generalizability of machine learning prediction models on attrition in longitudinal studies. *International Journal of Behavioral Development*, *46*(2), pp.169–176. https://doi.org/10.1177/01650254221075034

Jin, F.J.-Y., Maheshi, B., Lai, W., Li, Y., Gasevic, D., Chen, G., Charwat, N., Chan, P.W.K., Martinez-Maldonado, R., Gašević, D. and Tsai, Y.-S. (2025). Students' perceptions of generative AI-powered learning analytics in the feedback process: a feedback literacy perspective. *Journal of Learning Analytics*, *12*(1), pp.152–168. https://doi.org/10.18608/jla.2025.8609

Khosravi, H., Buckingham Shum, S., Chen, G., Conati, C., Tsai, Y.-S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S. and Gašević, D. (2022). Explainable artificial intelligence in education. *Computers and Education: Artificial Intelligence*, *3*, 100074. https://doi.org/10.1016/j.caeai.2022.100074

Kusuma, J., Halim, K., Pranoto, E., Kanigoro, B. and Irwansyah, E. (2022). Automated essay scoring using machine learning. In: *Proceedings of the 2022 International Conference on Cybernetics and Intelligent Systems (ICORIS)*, pp.1–5. https://doi.org/10.1109/ICORIS56080.2022.10031338

Küchemann, S., Avila, K.E., Dinc, Y., Hortmann, C., Revenga, N., Ruf, V., Stausberg, N., Steinert, S., Fischer, F., Fischer, M., Kasneci, E., Kasneci, G., Kuhr, T., Kutyniok, G., Malone, S., Sailer, M., Schmidt, A., Stadler, M., Weller, J. and Kuhn, J. (2025). On opportunities and challenges of large multimodal foundation models in education. *npj Science of Learning*, *10*(1), 11. https://doi.org/10.1038/s41539-025-00301-w

Lawasi, M., Rohman, V. and Shoreamanis, M. (2024). The use of AI in improving students' critical thinking skills. *Proceedings Series on Social Sciences & Humanities*, *18*, pp.366–370. https://doi.org/10.30595/pssh.v18i.1279

Li, H. and Botelho, A.F. (2024). Developing explainable AI systems to support feedback for students. In: *Proceedings of the 17th International Conference on Educational Data Mining*, pp.998–1002. https://doi.org/10.5281/zenodo.12730029

Mazzone, M. and Elgammal, A. (2019). Art, creativity, and the potential of artificial intelligence. *Arts*, *8*(1), 26. https://doi.org/10.3390/arts8010026

Messer, U. (2024). Co-creating art with generative artificial intelligence: Implications for artworks and artists. *Computers in Human Behavior: Artificial Humans*, *2*(1), 100056. https://doi.org/10.1016/j.chbah.2024.100056

Miao, F. and Cukurova, M. (2024). *AI competency framework for teachers*. Paris: UNESCO. https://doi.org/10.54675/ZJTE2084

Misgna, H., On, B.-W., Lee, I. and Choi, G.S. (2024). A survey on deep learning-based automated essay scoring and feedback generation. *Artificial Intelligence Review*, *58*(2), 36. https://doi.org/10.1007/s10462-024-11017-5

Nazaretsky, T., Mejia-Domenzain, P., Swamy, V., Frej, J. and Käser, T. (2025). The critical role of trust in adopting AI-powered educational technology for learning: an instrument for measuring student perceptions. *Computers and Education: Artificial Intelligence*, *8*, 100368. https://doi.org/10.1016/j.caeai.2025.100368

Ng, D.T.K., Leung, J.K.L., Su, J., Ng, R.C.W. and Chu, S.K.W. (2023). Teachers' AI digital competencies and twenty-first century skills in the post-pandemic world. *Educational Technology Research and Development*, *71*(1), pp.137–161. https://doi.org/10.1007/s11423-023-10203-6

Patterson, J.D., Barbot, B., Lloyd-Cox, J. and Beaty, R.E. (2024). AuDrA: An automated drawing assessment platform for evaluating creativity. *Behavior Research Methods*, *56*(4), pp.3619–3636. https://doi.org/10.3758/s13428-023-02258-3

Rachha, A. and Seyam, M. (2023). Explainable AI in education: current trends, challenges, and opportunities. In: *2023 IEEE SoutheastCon*, pp.232–239. https://doi.org/10.1109/SoutheastCon51012.2023.10115140

Sakhipov, A., Omirzak, I. and Fedenko, A. (2025). Beyond face recognition: a multi-layered approach to academic integrity in online exams. *Electronic Journal of e-Learning*, *23*(1), pp.81–95. https://doi.org/10.34190/ejel.23.1.3896

Samuel, K. (2024). The role of artificial intelligence in educational assessment. *Eurasian Experiment Journal of Scientific and Applied Research*, *5*, pp.44–48.

Saqr, M. and López-Pernas, S. (2024). Why explainable AI may not be enough: predictions and mispredictions in decision making in education. *Smart Learning Environments*, *11*(1), 52. https://doi.org/10.1186/s40561-024-00343-4

Spee, B.T.M., Mikuni, J., Leder, H., Scharnowski, F., Pelowski, M. and Steyrl, D. (2023). Machine learning revealed symbolism, emotionality, and imaginativeness as primary predictors of creativity evaluations of western art paintings. *Scientific Reports*, *13*(1), 12966. https://doi.org/10.1038/s41598-023-39865-1

Tenberga, I. and Daniela, L. (2024). Artificial intelligence literacy competencies for teachers through self-assessment tools. *Sustainability*, *16*(23), 10386. https://doi.org/10.3390/su162310386

Tierney, A., Peasey, P. and Gould, J. (2025). Student perceptions on the impact of AI on their teaching and learning experiences in higher education. *Research and Practice in Technology Enhanced Learning*, *20*, 005. https://doi.org/10.58459/rptel.2025.20005

U.S. Department of Education, Office of Educational Technology (2023). *Artificial intelligence and the future of teaching and learning: insights and recommendations*. Washington, DC.

Xu, W., Kassim, M.S.S., Hoo, W.L., Yang, W. and Xu, T. (2025). Explainable AI for education: Enhancing essay scoring via rubric-aligned chain-of-thought prompting. *International Journal of Modern Physics C, 37*(6), 2542013. https://doi.org/10.1142/S0129183125420136

Zhang, R., Lee, H., Wu, R., Yang, W. and Pan, Y. (2025). Is the impact of artificial intelligence generation tools on improving art education positive or negative? Perspectives of professors and students. *Interactive Learning Environments*, *33*(10), 5944–5975. https://doi.org/10.1080/10494820.2025.2490173

Zheng, C., Yu, Z., Jiang, Y., Zhang, M., Lu, X., Jin, J. and Gao, L. (2025). ArtMentor: AI-assisted evaluation of artworks to explore multimodal large language models capabilities. In: *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 659. https://doi.org/10.1145/3706598.3713274

From Intention to Reflection: Understanding Self-Directed Learning in the Use of Generative AI in Vietnam

Diep-Ngoc Le, Trinh Thi Nguyen, Huong-Giang Thi Doan, An Dieu Trinh and Mai-Huong Dinh Pham

VNU University of Education, Vietnam National University, Hanoi, Vietnam

lediep@vnu.edu.vn (corresponding author)
chingng103@gmail.com
hyang288@gmail.com
andieutrinh@gmail.com
phamh7027@gmail.com

https://doi.org/10.34190/ejel.24.2.4505

Abstract: This study extends the Theory of Planned Behavior (TPB) to explore how students' behavioral intentions toward using generative artificial intelligence (GenAI) are associated with their reflective engagement and self-directed learning (SDL) in higher education. As GenAI tools such as ChatGPT increasingly mediate learning, understanding how learners' intentions are linked to autonomous and reflective learning behaviors becomes essential. Data were collected from 149 first-year university students (predominantly female) in Vietnam who had prior experience with GenAI for academic purposes. Using Partial Least Squares Structural Equation Modeling (PLS-SEM), the study examined relationships among attitudes, subjective norms, perceived behavioral control, behavioral intention, actual use, reflection, and two dimensions of SDL, including intentional learning and self-management. The results reveal that attitudes and perceived behavioral control significantly predict students' intentions and actual use of GenAI, whereas subjective norms have no significant effect. Behavioral engagement is positively associated with reflection and both dimensions of SDL, while reflection is positively related to intentional learning and self-management, confirming its mediating role within the proposed model linking motivation-related constructs with autonomous learning outcomes. These findings highlight reflection as a metacognitive mechanism that links students' behavioral engagement with GenAI and their SDL-related outcomes. Theoretically, the study advances TPB by positioning reflection and SDL as outcome constructs within the proposed model, rather than fixed learner traits. Practically, it suggests that educators and institutions working with first-year university students or similar learner populations should integrate reflective activities and AI literacy into curricula to promote critical, ethical, and autonomous engagement with GenAI. Designing learning environments that position AI as a reflective partner, rather than merely a content generator, supports learners' self-regulation and reflective engagement. Overall, this research contributes to understanding how intentional and reflective interaction with GenAI is associated with deeper and more autonomous learning of students among first-year university students in a GenAI-supported learning context.

Keywords: Artificial intelligence, Generative AI, Theory of planned behavior, Self-directed learning, Reflective thinking, Higher education, PLS-SEM

1. Introduction

The rapid development of artificial intelligence (AI) has reshaped the higher education landscape, introducing new modes of teaching, learning, and assessment that transcend traditional classroom boundaries (S. Wang et al., 2024; Zawacki-Richter et al., 2019). Among these technologies, generative AI (GenAI) applications such as ChatGPT have emerged as transformative tools capable of generating adaptive and contextually relevant content, including text, images, and code (Abdallah et al., 2025; Li et al., 2025). These innovations enable personalized, interactive learning experiences that foster engagement and creativity, yet they also raise concerns about academic integrity and cognitive dependence (Banh & Strobel, 2023; Darwin et al., 2024). Consequently, understanding how university students regulate and reflect upon their learning when using GenAI has become an issue in higher education.

To explain students' behavioral patterns toward using GenAI, the Theory of Planned Behavior (TPB) has provided a valuable theoretical foundation (Ajzen, 1991; Ajzen, 2002). It posits that attitudes (ATT), subjective norms (SN), and perceived behavioral control (PBC) shape behavioral intention (INT) and subsequent action (BHE). TPB has been widely validated in technology acceptance studies, including mobile learning and online education (Aldammagh, Abdeljawad, & Obaid, 2021; Cheng, 2019; Teo, 2012). Recent research further confirms its relevance to understanding GenAI use, where learners' attitudes and trust strongly influence their intention to engage with these technologies (Bonsu & Baffour-Koduah, 2023; C. Wang et al., 2024; Wu & Dong, 2025). GenAI

Cite this article: Le, D. et al. 2026. "From Intention to Reflection: Understanding Self-Directed Learning in the Use of Generative AI in Vietnam", *Electronic Journal of e-Learning*, 24(2), pp 17-31, https://doi.org/10.34190/ejel.24.2.4505

technologies offer highly personalized and adaptive learning experiences (Li et al., 2025), which suggests a growing need to understand how intention translates into self-regulated learning behaviors in GenAI-supported environments. Learners are now required not only to form positive intentions toward using GenAI but also to regulate and direct their learning within these personalized environments (Ouaazki et al., 2024).

In this regard, self-directed learning (SDL) has gained increasing attention as a learning approach aligned with the adaptive and individualized nature of GenAI-supported education. When combined with digital technologies, SDL enables students to plan and manage their learning more effectively (Sumuer, 2018; Timothy et al., 2010). Empirical evidence further shows that SDL with technology (SDLT) enhances students' readiness and predicts both their intention and actual use of AI-based chatbots in higher education (Esiyok, Sahin, & Kucukergin, 2024). Grounded in Knowles' (1975) concept that learners need to develop proactive and reflective skills with objects in the learning process. Recent studies have explored students' initial reflections during GenAI use (Šedlbauer et al., 2024) and examined how GenAI tools can scaffold reflective thinking (RT) (Wei et al., 2025); these studies have primarily viewed SDL as a pre-existing learner trait rather than a developmental outcome. This present study aims to investigate how students interact with GenAI tools by focusing on intention, actual behaviour use, reflective thinking engagement, and SDL experiences. Within this process, reflection serves as a metacognitive mechanism that allows learners to monitor, evaluate, and adapt their Learning with GenAI.

Complementing this perspective, reflection allows individuals to analyze experiences, connect ideas, and make informed judgments (Lee, 2005; Rodgers, 2002). The immediacy and personalization of feedback offered by GenAI provide unique opportunities for fostering reflective engagement in learning (Ficko et al., 2025; Saritepeci & Yildiz Durak, 2024). Despite growing recognition of this potential, empirical evidence on how RT interacts with SDL in GenAI-based learning remains limited. To address these gaps, the present study integrates TPB, SDL, and RT into a unified framework to examine how university students adopt and learn with GenAI tools. Specifically, it investigates how behavioral intentions toward GenAI are transformed into actual learning behaviors and reflective processes that, in turn, enhance SDL outcomes. By empirically validating this extended model using PLS-SEM, the study aims to deepen theoretical understanding and offer practical insights into promoting reflective and SDL in using GenAI.

2. Literature Review

2.1 Theory of Planned Behavior with using Generative AI

The Theory of Planned Behavior (TPB) was proposed by Ajzen (1991), a theoretical model designed to explain and predict human behavior. TPB assumes that human social behavior is largely under volitional control and can be predicted from an individual's intentions. These intentions are driven by three principal factors: attitude (ATT), subjective norms (SN), and perceived behavioral control (PBC). Ajzen (1991) further divided PBC into controllability and self-efficacy, with facilitating conditions often employed to represent the concept of PBC. Within TPB, behavioral intention (INT) is considered the most influential predictor of actual behavior.

The TPB has been widely applied in studies on technology acceptance (Cheng, 2019; Teo, 2012; C. Wang et al., 2024), including research on electronic recruitment (Parikh, Patel, & Jaiswal, 2021), mobile banking (Aldammagh, Abdeljawad, & Obaid, 2021), and mobile learning (Gómez-Ramirez, Valencia-Arias, & Duque, 2019). A study by Teo (2012) on pre-service teachers' intentions to use technology revealed that their attitude toward computer use had the greatest influence on intention; SN and facilitating conditions (representing PBC) also affected behavioral intention. Consistent with this, Venkatesh, Thong, & Xu (2012) empirically confirmed that intention (INT) is a strong predictor of actual technology usage behavior (BHE). Notably, TPB offers a more comprehensive account of intention and behavior in online collaborative learning, highlighting the combined effects of social and control factors (Cheng, 2019).

In the context of AI, GenAI, and ChatGPT, the TPB and its integrated models have remained appropriate theoretical frameworks for examining technology acceptance and use (Bonsu & Baffour-Koduah, 2023; Phuong Dung et al., 2023). GenAI tools such as ChatGPT, with their ability to generate novel and realistic content such as text and images (Banh & Strobel, 2023), have transformed the way learners interact with information and engage in learning processes (Bozkurt, 2023; B. Li et al., 2024). The core constructs of TPB have continued to be investigated: learners' attitudes and trust in ChatGPT as a learning tool have played a crucial role in shaping their self-directed language learning activities (Bearman & Ajjawi, 2023). The SN has primarily exerted its influence through the formation of a positive attitude toward GenAI, which subsequently affects usage intention. In the context of GenAI use, ATT has served as the key determinant of behavioral intention, while SN and PBC played indirect or supporting roles (C. Wang et al., 2024). The study by Wu and Dong (2025) has demonstrated that SN is a key factor exerting a direct and significant influence on the intention to use GenAI. In summary, the TPB has remained a powerful theoretical model for understanding technology acceptance intentions and behaviors, including the rapidly evolving GenAI tools. Although the TPB has been extensively validated, applying it to GenAI use calls for a deeper inquiry into how learners' behavioral intentions give rise to autonomous learning behaviors promoting SDL, thereby reducing potential overreliance on AI tools.

H_1: ATT will have a positive influence on INT.

H_2: SN will have a positive influence on INT.

H_3: PBC will have a positive influence on INT.

H_4: INT will have a positive influence on BHE.

2.2 Self-Directed Learning with Technologies

The concept and role of self-directed learning (SDL) were articulated by the renowned American adult education scholar Malcolm Knowles, who defined it as follows: The SDL refers to a process in which individuals take the initiative to identify their learning needs, set goals, select strategies, and evaluate outcomes, with or without assistance from others (Knowles, 1975, p. 18). For each learner, SDL has occurred through a process involving personalized guidance within the broader teaching and learning process (Hiemstra, 1988). SDL has also been viewed as an approach in which learners assume personal responsibility for self-monitoring cognition and self-managing the learning context to achieve meaningful outcomes (Garrison, 1997). In the modern context, SDL has enabled learners to develop greater control over knowledge construction, engage in responsible self-reflection regarding their learning and learning strategies, and enhance their performance through self-evaluation of self-learning activities (Gibbons, 2002).

In higher education, SDL has become increasingly important in equipping students with the essential skills and competencies required to succeed in the digital era and to engage in lifelong learning (Boyer et al., 2013). Recent work on AI-enhanced information and communication technologies in lifelong learning suggests that AI can be embedded in person-oriented learning environments that support continuous learning and professional retraining through adaptive and personalized opportunities (Papadakis et al., 2024). In addition, qualitative evidence on SDL supported by the use of ChatGPT suggests that learners may incorporate GenAI into planning, monitoring, and evaluating their learning, while the broader socio-technical landscape of GenAI continues to shape the learning context in which SDL unfolds (B. Li et al., 2024). Practicing SDL has served as a vital instrument for personal development in the twenty-first century, enabling students to enhance the quality of their learning and prepare effectively for future careers (Boyer et al., 2013; Dehnad et al., 2014). The advancement of information and communication technology (ICT) has created a favorable environment for students to strengthen their engagement in SDL. The use of technological tools has enabled students to access a vast array of diverse information resources, locate and evaluate information, monitor the process of planning and implementation, and subsequently self-assess their learning progress (Lee et al., 2014). Students' online learning outcomes are indirectly influenced by their online learning readiness, which is mediated by individual competencies such as SDL skills, metacognitive ability, and collaborative skills (Ho, Kuo, & Lin, 2009).

Research on developing SDL measurement scales for university students has attracted growing attention. Hendry and Ginns (2009) have examined SDL as a key construct influencing learners' ability to plan, manage, and evaluate their own learning. In the context of a private university in Singapore, Khiat (2017) identified various indicators of SDL, such as goal setting, time management, preparation, and readiness for learning, and found that they had both direct and indirect effects on students' academic performance. Additionally, Timothy et al. (2010) developed the SDL with Technology Scale (SDLTS) and identified two key components: self-management (SDL_M), which reflects learners' ability to plan and regulate their learning processes, and intentional learning (SDL_I), which represents learners' goal-oriented pursuit of knowledge, encompassing the identification of both academic and non-academic learning goals and the application of acquired skills and knowledge to new contexts. Subsequently, employing a modified version of the SDLTS, Sumuer (2018) has shown that computer self-efficacy had a significant and indirect impact on university students' engagement in SDLT.

The emergence of AI, GenAI tools like ChatGPT, has significantly transformed the landscape of SDL among learners (B. Li et al., 2024; Ouaazki et al., 2024). GenAI has been regarded as a tool capable of revolutionizing the SDL process by providing unprecedented opportunities for personalized education and adaptive learning (Wu et al., 2024). ChatGPT has held great potential to support SDL by offering instant feedback and engaging in interactive communication tailored to learners' individual characteristics and ongoing learning processes (B. Li et al., 2024). ChatGPT has served as a virtual tutor (Lin, 2024), an AI proxy teacher (Chiu et al., 2024), or a personalized learning assistant (Wu et al., 2024) that helps students enhance their learning experience. Personalized learning has been supported by ChatGPT can foster SDL behaviors among students and lead to improved learning outcomes (Li et al., 2025). In this regard, learners' behavioral engagement with GenAI represents a key driver that facilitates both intentional learning and self-management dimensions of SDL

H_5: BHE will have a positive influence on SDL_I.

H_6: BHE will have a positive influence on SDL_M.

2.3 Reflections on Generative AI

The fast-paced evolution of AI has brought about a revolution in the field of education (Abdallah et al., 2025; Chen, Chen, & Lin, 2020; S. Wang et al., 2024; Zawacki-Richter et al., 2019). Particularly with the emergence of GenAI and applications such as ChatGPT, which are capable of generating new and realistic content, including text, images, and programming code, based on users' basic prompts (Banh & Strobel, 2023; Bordas et al., 2024). GenAI has offered numerous significant benefits for higher education by providing personalization and adaptivity in learning (Adiguzel, Kaya, & Cansu, 2023; Atlas, 2023; Li et al., 2025), enabling the customization of curricula and learning materials according to individual students' needs (B. Li et al., 2024), thereby enhancing the overall learning experience and educational quality (Chen, Chen, & Lin, 2020). GenAI tools have assisted instructors in designing differentiated lesson plans, creating personalized assignments, and analyzing or assessing students' competencies (Monzon & Hays, 2025). For students, GenAI has facilitated interactive learning, instant feedback (Chang et al., 2024), knowledge exploration, problem-solving (Ouaazki et al., 2024), and academic writing (Wang, Li, & Bonk, 2024). Moreover, GenAI has contributed to developing learning habits, fostering interest and motivation (Li et al., 2025; Monzon & Hays, 2025), particularly in the domain of SDL (Bosch & Kruger, 2024; Esiyok, Sahin, & Kucukergin, 2024; Belle Li et al., 2024).

However, the integration of GenAI into higher education has also presented numerous challenges and risks (Banh & Strobel, 2023). One of the major concerns has been the potential increase in students' dependence and complacency (Ouaazki et al., 2024). Students have been overusing GenAI to complete assignments quickly, leading to superficial learning and a decline in critical thinking ability (Darwin et al., 2024; van den Berg & du Plessis, 2023). The accuracy and reliability of AI-generated content have also been worrisome, as GenAI has produced information that appears plausible but is actually incorrect or fabricated (Banh & Strobel, 2023) or delivers biased responses, necessitating significant effort for verification (B. Li et al., 2024). Other ethical concerns include risks related to privacy, copyright infringement, and the lack of transparency in GenAI's data collection and processing mechanisms (Banh & Strobel, 2023). Given these challenges, fostering RT becomes essential for ensuring that learners engage with GenAI critically and responsibly. At the same time, recent pedagogical scholarship argues that higher education should prepare students to engage with AI systems whose internal decision-making processes are not fully transparent or traceable, emphasizing the development of evaluative judgement through orientation to quality standards and meaningful interactions with AI systems (Bearman & Ajjawi, 2023). In this framing, AI tools can be positioned less as shortcut mechanisms and more as learning partners that support lifelong learning orientations when their use is pedagogically guided, ethically grounded, and aligned with SDL processes (Papadakis et al., 2024; B. Li et al., 2024).

In this context, the RT has emerged as a crucial construct for mediating how learners engage with GenAI responsibly and meaningfully (Šedlbauer et al., 2024; Wei et al., 2025). Reflection is a process of constructing meaning and enables learners to move from one experience to another with a deeper understanding of how these experiences and ideas are interconnected (Rodgers, 2002). RT enables learners in higher education to critically examine their learning and professional actions, fostering continuous improvement and greater insight (Lee, 2005). A four-level instrument construct has been applied to measure students' RT, including: habitual action, understanding, reflection, and critical reflection (Kember et al., 2000). In addition, the frequency of RT occurring and the objects that learners reflect on have also been utilized to evaluate students' RT (Hong & Choi, 2015).

Further, Rodgers (2002) has stated that the manifestations of reflection will appear through interaction and dialogue with others. Hence, the personalized nature of the interaction and the near-immediate feedback provided by GenAI provided opportunities that are suitable for performing RT by the learner during SDL (Ficko et al., 2025; Wu et al., 2024). Recent research has emphasized the potential of GenAI to trigger reflection (Lee, 2005; Wei et al., 2025), but the mechanisms through which such reflection contributes to SDL need to be explored. This gap highlights the need for empirical research to examine how RT mediates or enhances SDL when learners interact with GenAI tools.

H_7: BHE will have a positive influence on RT.

H_8: RT will have a positive influence on SDL_I.

H_9: RT will have a positive influence on SDL_M.

The study proposes nine hypotheses to examine the structural relationships among the constructs. Specifically, ATT, SN, and PBC are hypothesized to influence INT (H_1-H_3). The INT is expected to affect BHE and SDL (H_4), while BHE is assumed to impact RT, SDL_I, and SDL_M (H_5-H_7). Finally, RT is hypothesized to influence SDL_I and SDL_M (H_8-H_9). Overall, the proposed model extends TPB by incorporating RT as a cognitive mechanism through which SDL_I and SDL_M. The proposed research model is presented in Figure 1.

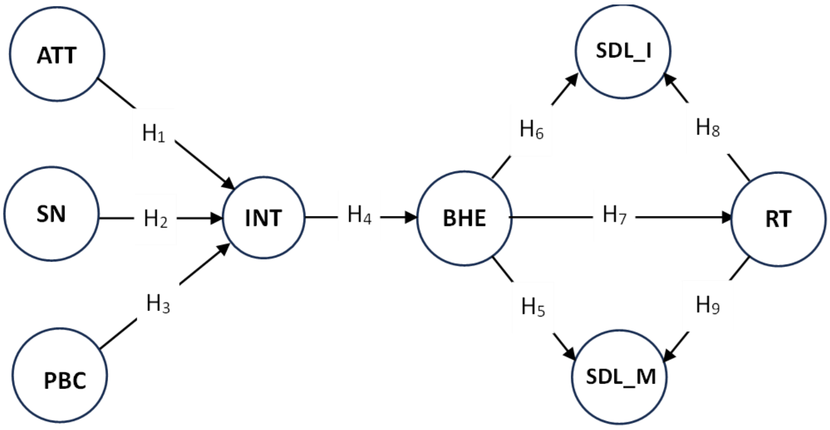


Figure 1: Proposed Research Model

3. Research Methodology

A total of 149 students participated in the survey. In terms of gender, females accounted for the majority with 91.3% (136 students), while males represented 8.7% (13 students). All participants were first-year students, and the survey link was distributed through first-year class groups, allowing students from different classes to voluntarily participate based on their prior experience using GenAI in their learning processes (Table 1). While this sampling approach was appropriate for the study context, the resulting gender imbalance constitutes a methodological limitation that should be taken into account when interpreting the findings.

Table 1: Demographic Distribution of Participants

Measure	Item	N	Percentage
Gender	Male	13	8.7%
	Female	136	91.3%
Which GenAI application have you used the most for your learning?	ChatGPT	106	71.1%
	Gemini	27	18.1%
	Microsoft Copilot	8	5.4%
	DeepSeek	8	5.4%

The study adopted a quantitative research approach, employing a cross-sectional survey design to investigate the relationships among the TPB, SDL, and RT in the context of using GenAI for learning. Data were gathered through a 24-item questionnaire, which incorporated measurement scales adapted from prior studies and refined to align with the context of Vietnamese university students. The instrument consisted of statements rated on a five-point Likert scale (ranging from “strongly disagree” to “strongly agree”) and items modified from existing validated scales. The questionnaire was translated into Vietnamese, and language accuracy was ensured through expert consultation and revision (see Table 2).

Table 2: Descriptive Statistics

Construct	Items	M	SD	Sources
Attitude	ATT1. I think using GenAI (like ChatGPT) is useful for my learning.	3.899	0.939	(Ajzen, 1991; Teo, 2012; Teo & van Schaik, 2012)
	ATT2. I think using GenAI will help me study more effectively.	3.906	0.763	
	ATT3. I think using GenAI for learning is a positive thing and should be encouraged.	3.772	0.844	
Subjective Norm	SN1. My friends think that I should use GenAI to support my learning.	3.839	0.751	
	SN2. I feel influenced by people around me in my decision to use GenAI.	3.711	0.780	
	SN3. My lecturers or university support my use of GenAI in learning.	3.691	0.759	
Perceived Behavioral Control	PBC1. I have enough skills to use GenAI tools for my studies.	3.872	0.735	
	PBC2. I know how to control my use of GenAI to match my learning goals.	3.906	0.679	
	PBC3. Even without guidance, I feel confident using GenAI tools effectively.	3.705	0.799	
Intention	INT1. I plan to use (or continue using) GenAI tools to support my learning in the future.	4.000	0.769	
	INT2. I want to explore different ways to use GenAI to improve my learning.	4.128	0.744	
	INT3. I am willing to try different GenAI tools to make my learning more effective.	4.081	0.790	
Behavior	BHE1. I often use GenAI tools to complete my study tasks.	3.899	0.801	(Venkatesh, Thong, & Xu, 2012).
	BHE2. I have shared GenAI responses or content with my classmates for discussion and learning.	3.899	0.792	
	BHE3. I regularly use GenAI tools for my study and research activities.	4.000	0.695	
Reflective Thinking	RT1. I think using GenAI helps me learn better, but it cannot replace personal effort.	4.007	0.798	(Kember et al., 2000; Šedlbauer et al., 2024)
	RT2. I feel responsible for using AI tools ethically in my learning.	4.081	0.755	
	RT3. I think I need to learn more about how to check and evaluate the content created by GenAI.	4.060	0.813	
Intentional Learning	SDL_I1. I look for more information using GenAI to help me understand my lessons better.	3.953	0.679	(Timothy et al., 2010)
	SDL_I2. I use GenAI tools to find and organize information that supports my learning goals.	3.913	0.759	
	SDL_I3. I use GenAI to become better at a skill I am interested in.	3.805	0.792	
	SDL_I4. I use GenAI to get ideas from websites and other people to learn more about a topic.	3.899	0.801	
Self-Management	SDL_M1. I use GenAI to ask questions about my lessons when I don't understand something, so I can learn by myself.	4.054	0.703	
	SDL_M2. I use GenAI to express and develop my own thoughts and ideas for school tasks or assignments.	3.893	0.725	

The participants of the survey were undergraduate students from a university in Vietnam who had prior experience with or interest in utilizing GenAI tools for learning purposes. Data were collected online via Google Forms during May 2025. After screening and data cleaning, a total of 149 valid responses were retained for analysis. The sample size was deemed adequate for Partial Least Squares Structural Equation Modeling (PLS-SEM) analysis based on the criteria suggested by Hair et al. (2019). Firstly, given the study parameters (significance level $\alpha = 0.05$, statistical power = 0.90, and effect size $f^2 = 0.15$), and according to the research model (Figure 1), the maximum number of independent variables influencing a single dependent variable was three. The actual sample size was 149, and this configuration aligns with the sample size of 88 when estimation was obtained using G*Power (Kock & Hadaya, 2018). Secondly, the obtained sample met the minimum thresholds recommended by Hair et al. (2010) and Hair et al. (2019). The PLS-SEM procedure followed two main stages: assessment of the measurement model and evaluation of the structural model. Firstly, for the measurement model, item reliability was verified with factor loadings expected to exceed 0.70. Construct reliability was assessed using Cronbach's alpha and composite reliability (CR), both with a recommended minimum threshold of 0.70 (Hair et al., 2019). Convergent validity was established when the average variance extracted (AVE) was greater than 0.50. Discriminant validity was confirmed through the Fornell-Larcker criterion and the heterotrait-monotrait ratio (HTMT), both expected to be below 0.85 (Fornell & Larcker, 1981; Henseler, Ringle, & Sarstedt, 2015). For the structural model, explanatory power was evaluated based on the R^2 values (0.25 = weak, 0.50 = moderate, 0.75 = substantial) (Hair et al., 2019), effect sizes f^2 (0.02 = small, 0.15 = medium, 0.35 = large), and predictive relevance Q^2 (> 0). Finally, bootstrapping with 5,000 resamples was performed to test the significance of the path coefficients ($t > 1.96$, $p < 0.05$) (Hair et al., 2019).

4. Results

4.1 Measurement Model

The factor loadings in this study ranged from 0.706 to 0.925 (see Table 3), most of which met the recommended threshold suggested by Hair et al. (2019). SN3 demonstrated a loading of 0.672, which is slightly below the recommended threshold of 0.70. It was retained because the overall reliability (CR = 0.858) and convergent validity (AVE = 0.671) of the SN construct remained acceptable. To ensure convergent validity, both the composite reliability (CR) and average variance extracted (AVE) values must be at least 0.50 (Hair et al., 2019). In this model, the CR values ranged from 0.848 to 0.914, and the AVE values ranged from 0.651 to 0.841 (see Table 3), confirming that the measurement model achieved satisfactory levels of reliability and convergent validity. Furthermore, the Cronbach's Alpha and rho_A values for all constructs exceeded 0.70, indicating satisfactory internal consistency reliability.

The overall model fit was assessed using SRMR and NFI. The SRMR value of the saturated model was 0.079, confirming good model fit, while the NFI value of 0.700 was within the acceptable range for PLS-SEM (Hair et al., 2019). Collinearity was assessed through the variance inflation factor (VIF), and all indicators showed acceptable VIF values between 1.215 and 2.295, well below the conservative cut-off value of 3.3 (Hair et al., 2021; Kock, 2017), indicating no multicollinearity issue.

Next, the study examined discriminant validity using both the Fornell–Larcker and Heterotrait-Monotrait ratio (HTMT) criteria. The square roots of the AVE values (displayed on the diagonal of Table 4) were greater than all corresponding inter-construct correlations, confirming discriminant validity following (Fornell & Larcker, 1981). Furthermore, as recommended by Franke and Sarstedt (2019), most HTMT values were below the 0.90 threshold, except for the ATT-SN pair (HTMT = 0.972), which slightly exceeded the cut-off. However, given that the constructs remain theoretically distinct and other validity criteria (Fornell-Larcker and cross-loadings) were satisfied, the measurement model still demonstrates acceptable standards for validity and reliability (Hair et al., 2019).

Table 3: Reliability and Convergent Validity Indices of the Measurement Scale

Components	Outer loading	VIF	Cronbach's Alpha	rho_A	CR	AVE
ATT1	0.752	1.388	0.781	0.805	0.873	0.697
ATT2	0.898	2.047				
ATT3	0.849	1.885				
BHE1	0.821	1.725	0.818	0.831	0.892	0.734
BHE2	0.838	1.800				
BHE3	0.908	2.275				
INT1	0.822	1.701	0.803	0.806	0.884	0.719
INT2	0.893	2.145				
INT3	0.826	1.659				
PBC1	0.706	1.215	0.727	0.735	0.848	0.651
PBC2	0.842	1.778				
PBC3	0.864	1.862				
RT1	0.866	2.013	0.849	0.850	0.909	0.768
RT2	0.892	2.295				

Components	Outer loading	VIF	Cronbach's Alpha	rho_A	CR	AVE
RT3	0.871	1.973				
SDL_I1	0.776	1.532	0.834	0.836	0.889	0.668
SDL_I2	0.841	1.956				
SDL_I4	0.824	1.910				
SDL_I3	0.827	2.017				
SDL_M1	0.925	1.876	0.812	0.816	0.914	0.841
SDL_M2	0.910	1.876				
SN1	0.910	1.749	0.770	0.883	0.858	0.671
SN2	0.856	1.688				
SN3	0.672	1.421				

*Note.*VIF = Variance Inflation Factor; rho_A = construct reliability; CR = composite reliability; AVE = average variance extracted; ATT = Attitude; SN = Subjective Norm; PBC = Perceived Behavioral Control; INT = Intention; BHE = Behavior; SDL_I = Intentional Learning; SDL_M = Self-Management; RT = Reflective Thinking.

4.2 Structural Model

The hypotheses of the model were tested using the bootstrapping method (Hair et al., 2021; Hair et al., 2019). The results indicated that ATT had a strong and significant positive effect on INT ($\beta = 0.557$; $t = 5.063$; $p < 0.001$), while PBC also significantly influences INT ($\beta = 0.339$; $t = 3.945$; $p < 0.001$). However, the effect of SN on INT was not significant ($\beta = -0.035$; $t = 0.288$; $p = 0.778$). Moreover, INT showed a strong association with BHE ($\beta = 0.782$; $t = 12.009$; $p < .001$) (see Table 4).

In addition, BHE significantly affected SDL_I ($\beta = 0.413$; $t = 5.683$; $p < 0.001$), SDL_M ($\beta = 0.289$; $t = 2.728$; $p < 0.005$), and RT ($\beta = 0.550$; $t = 5.344$; $p < 0.001$). Similarly, RT positively influences both SDL_I ($\beta = 0.473$; $t = 5.303$; $p < 0.001$) and SDL_M ($\beta = 0.556$; $t = 5.945$; $p < 0.001$). These findings confirmed that most of the direct hypotheses in the research model are statistically supported (see Fig. 2).

Regarding the indirect effects, the results showed that ATT and PBC had significant indirect influences on SDL_I and SDL_M through INT to BHE and BHE to RT pathways. The indirect effect of ATT on SDL_I and SDL_M through INT to BHE to RT was approximately $\beta = 0.113$ and 0.133. Similarly, PBC indirectly influences SDL_I and SDL_M ($\beta = 0.069$ and $\beta = 0.081$). However, SN was not statistically supported to impact learners' SDL indirectly. In short, these indirect effects were positive and meaningful, demonstrating that RT serves as the key mediating mechanism construct within the proposed model linking individuals' intentions and behavioral outcomes (see Table 5).

Table 4: Fornell-Larcker and HTMT Indices for the Discriminant Validity

Fornell-Larcker Criterion	ATT	BHE	INT	PBC	RT	SDL_I	SDL_M	SN
ATT	0.835							
BHE	0.687	0.857						
INT	0.677	0.782	0.848					
PBC	0.438	0.552	0.565	0.807				
RT	0.510	0.550	0.651	0.396	0.877			
SDL_I	0.606	0.673	0.678	0.418	0.700	0.817		
SDL_M	0.576	0.594	0.580	0.510	0.714	0.803	0.917	
SN	0.789	0.650	0.572	0.495	0.384	0.540	0.464	0.819
Heterotrait-Monotrait Ratio								
ATT								
BHE	0.853							
INT	0.848	0.959						
PBC	0.592	0.719	0.740					

Fornell-Larcker Criterion	ATT	BHE	INT	PBC	RT	SDL_I	SDL_M	SN
RT	0.614	0.654	0.787	0.509				
SDL_I	0.738	0.808	0.822	0.656	0.825			
SDL_M	0.708	0.727	0.716	0.550	0.858	0.970		
SN	0.972	0.797	0.652	0.680	0.420	0.606	0.550	

Note. ATT = Attitude; SN = Subjective Norm; PBC = Perceived Behavioral Control; INT = Intention; BHE = Behavior; SDL_I = Intentional Learning; SDL_M = Self-Management; RT = Reflective Thinking.

Table 5: Path Coefficients of the Structural Model

Relationship	β	SD	t-value	f^2	Result
ATT → INT	0.557	0.110	5.063**	0.257	Supported
BHE → RT	0.550	0.103	5.344**	0.433	Supported
BHE → SDL_I	0.413	0.073	5.683**	0.305	Supported
BHE → SDL_M	0.289	0.106	2.728**	0.135	Supported
INT → BHE	0.782	0.065	12.009**	1.578	Supported
PBC → INT	0.339	0.086	3.945**	0.190	Supported
RT → SDL_I	0.473	0.089	5.303**	0.400	Supported
RT → SDL_M	0.556	0.093	5.945**	0.500	Supported
SN → INT	- 0.035	0.123	*0.288*	0.001	Unsupported
Specific Indirect Effects					
ATT → INT → BHE → RT → SDL_I	0.113	0.045	2.540*		Supported
PBC → INT → BHE → RT → SDL_I	0.069	0.031	2.209*		Supported
SN → INT → BHE → RT → SDL_I	- 0.007	0.027	*0.270*		Unsupported
ATT → INT → BHE → RT → SDL_M	0.133	0.053	2.534*		Supported
PBC → INT → BHE → RT → SDL_M	0.081	0.036	2.252*		Supported
SN → INT → BHE → RT → SDL_M	- 0.008	0.031	*0.275*		Unsupported

Note: SD = Standard Deviation; ATT = Attitude; SN = Subjective Norm; PBC = Perceived Behavioral Control; INT = Intention; BHE = Behavior; SDL_I = Intentional Learning; SDL_M = Self-Management; RT = Reflective Thinking.

Significance levels: * p < 0,05; ** p < 0,01

4.3 Explanatory Power and Effect Sizes

The analysis results demonstrate that the extended model provides a solid explanation for behavioral and learning outcomes among university students. Specifically, the R^2 value of INT reached 0.539, indicating a moderate level of explanatory power, while BHE (R^2 = 0.609) also showed a high level of explanatory power. In contrast, RT (R^2 = 0.297) demonstrates a lower but still meaningful level of explanatory power. The SDL_I (R^2 = 0.604) and SDL_M (R^2 = 0.563) achieved moderate explanatory power (see Table 6). These findings confirm that the proposed model accounts for a considerable proportion of the variance in the key endogenous constructs.

Regarding the effect sizes (f^2), the path from INT to BHE showed the strongest effect (f^2 = 1.578), representing a very large contribution to the explained variance. The effects of ATT on INT (f^2 = 0.257) and PBC on INT (f^2 = 0.190) were also substantial, while the SN on INT showed a negligible impact (f^2 = 0.001). Moreover, RT demonstrated a meaningful role within the model, exhibiting a strong effect on SDL_I (f^2 = 0.400) and an effect on SDL_M (f^2 = 0.500). Overall, these results highlight the central role of BHE and RT as a mediating construct that bridges intention and actual behavior, supporting the view that individuals' ability to regulate and direct their own learning is pivotal in translating motivational factors into effective behavioral outcomes. The combination of high R^2, significant effect sizes, and positive predictive relevance (Q^2 values ranging from 0.219 to 0.450) confirms that the structural model achieves both explanatory and predictive adequacy in this study context (see Table 6).

Table 6: Coefficient of Determination and Predictive Relevance of the Structural Model

Variable	R^2 Adjusted	Q^2
INT	0.539	0.376
BHE	0.609	0.436
RT	0.297	0.219
SDL_I	0.604	0.379
SDL_M	0.563	0.450

Note: INT = Intention; BHE = Behavior; SDL_I = Intentional Learning; SDL_M = Self-Management; RT = Reflective Thinking.

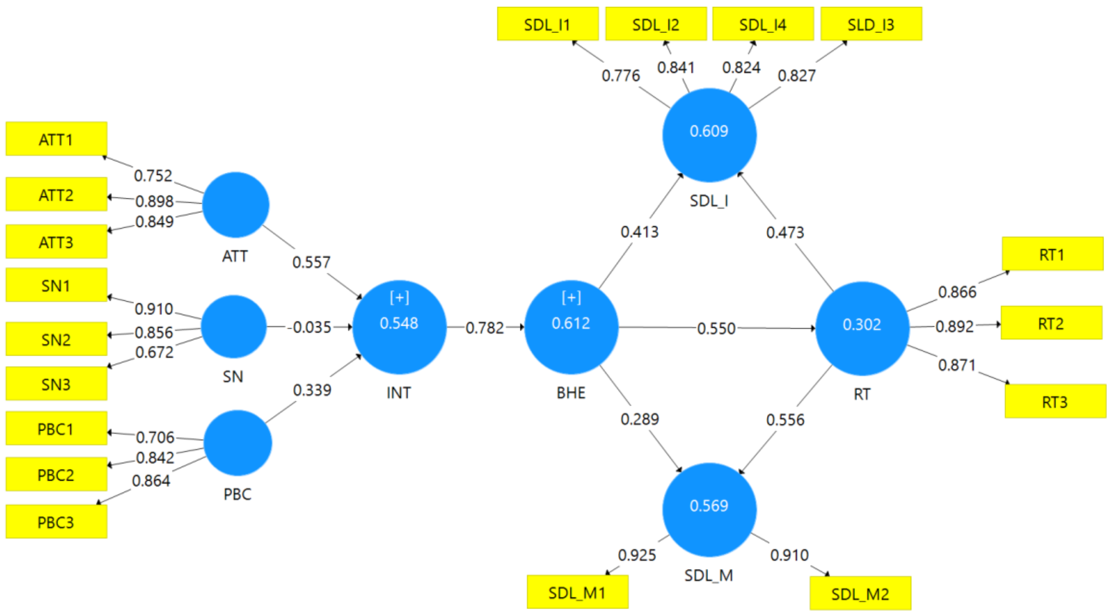


Figure 2: Empirical Model

5. Discussion

The findings of the present study suggest a sequential pattern of associations within the proposed model, whereby behavioral intention is linked to actual use, which is associated with reflective engagement and, in turn, with SDL dimensions. This pathway extends the TPB (Ajzen, 1991; Ajzen, 2002) by explaining how motivational factors are associated with autonomous learning behaviors through reflection within the proposed model. The non-significant effect of SN supports prior research (Cheng, 2019; Watson & Rockinson-Szapkiw, 2021; Yao et al., 2022). It suggests that social influence plays a limited role in AI-supported learning environments similar to the present study context, where self-direction and independence are more central (B. Li et al., 2024; Li et al., 2025). In GenAI settings, characterized by personalized and immediate feedback (Abdallah et al., 2025), learners' choices are shaped more by intrinsic motivation and self-regulation than by external expectations. From a gender perspective, prior studies have reported that differences in technology-supported learning extend beyond technology acceptance to self-regulatory processes central to SDL and reflective learning practices (Møgelvang et al., 2024; Sobieraj & Krämer, 2020). Female students, in particular, have been associated with greater emphasis on self-regulation underpinning autonomous and SDL, reflective engagement, and autonomous learning management in digital learning contexts (Møgelvang et al., 2024; Sobieraj & Krämer, 2020). Given that the present sample consists predominantly of female first-year students, peer influence may be particularly salient during the early stages of the university transition. However, within the scope of the proposed cross-sectional model, SDL-related behaviors are more closely associated with reflective engagement and self-regulatory mechanisms, while the influence of social norms may operate in more indirect or temporally diffuse ways that are less salient at a single point of measurement. This gender-skewed sample composition may therefore help explain why SN did not emerge as a significant predictor in the proposed model, despite their established role in technology acceptance research (Goswami & Dutta, 2016). In addition, it is possible that the non-significant effect of SN reflects differences in the sources of normative influence captured by the measurement, as perceptions of peer expectations and institutional or instructor expectations may exert opposing influences, potentially canceling each other within the SN construct. In addition to substantive explanations related to gender composition and learning context, the role of statistical power should also be considered when interpreting the non-significant paths involving SN. Although the sample size met the minimum requirements for PLS-SEM analysis (Hair et al., 2019), it may have limited the detection of smaller or unstable effects of SN that have been shown to vary across contexts in technology acceptance research (Venkatesh et al., 2003). Accordingly, the absence of statistically significant effects for SN should be interpreted with caution, as it may reflect limited power rather than the complete absence of social influence.

This research reconceptualizes SDL as an outcome construct within the proposed model, rather than solely as a fixed learner trait. In doing so, it addresses a theoretical gap in the current literature. Previous research often regarded SDL as an antecedent of technology adoption (Esiyok, Sahin, & Kucukergin, 2024; Sumuer, 2018; Timothy et al., 2010), treated SDL as an antecedent of technology adoption, whereas the current findings show that engagement with GenAI, when accompanied by reflective engagement, is associated with higher levels of SDL-related behaviors within the model. This perspective aligns with the view that learners can be strengthened through guided reflection and purposeful engagement (Lee et al., 2014), and supports recent empirical evidence that reflection is the link between AI interaction and meaningful learning (Šedlbauer et al., 2024). Whereas Wei et al. (2025) conceptualized RT as an outcome of GenAI-assisted learning, this study identifies RT as the underlying mechanism through which behavioral engagement is associated with SDL dimensions. It should be acknowledged that the observed associations do not establish the direction of influence between GenAI use and reflective engagement. It is possible that reflective learners engage with GenAI in more reflective ways, just as structured interaction with GenAI may afford opportunities for reflection. Longitudinal or experimental designs are required to disentangle these two-way relationships.

For practical and pedagogical implications, the results show the need to position GenAI as a reflective partner rather than a simple information provider. Within the scope of the observed associations, GenAI-based learning tasks may be designed to emphasize personal goal-setting and self-monitoring features that support learners' intrinsic motivation. The link between RT and SDL found in this study shows that integrating structured reflection activities can directly reinforce learners' autonomy in learning. Educators should design AI-integrated tasks that require learners to analyze, critique, and justify AI-generated outputs, as such reflective activities promote deeper engagement and self-regulation (Ficko et al., 2025; Saritepeci & Yildiz Durak, 2024). Curriculum designers and teacher educators should also integrate reflective learning tasks, such as AI–human comparison or justification writing, to reinforce the link between GenAI use and SDL (Li et al., 2025; Ouaazki et al., 2024). Furthermore, professional development for instructors should focus on cultivating students' reflective and autonomous skills when using GenAI responsibly, mitigating the risks of overreliance and superficial learning noted by Banh & Strobel (2023).

At the institutional and policy level, the study's findings support integrating AI literacy and reflective learning competencies into higher education frameworks (S. Wang et al., 2024; Zawacki-Richter et al., 2019). Universities should develop assessment systems that evaluate not only students' technical proficiency with GenAI but also their reflective and self-directed capacities. Institutional policies should emphasize transparency and ethical use of AI tools, encouraging students to engage critically and autonomously with technology while maintaining academic integrity (Abdallah et al., 2025). Additionally, institutional AI policies can promote structured reflection cycles within learning management systems or GenAI platforms, ensuring that learners engage in monitoring, feedback, and self-evaluation. These initiatives align with the global movement to foster autonomous, lifelong learners who can navigate AI-driven educational ecosystems responsibly.

Although the PLS-SEM results provide robust evidence for the proposed model, several limitations should be acknowledged. The cross-sectional design restricts causal inference; future longitudinal or experimental research could better capture the developmental nature of SDL across different stages of GenAI use. The sample was drawn from first-year students across multiple classes within a single university. Although participants represented diverse classes within the institution, the use of a single-institution sample limits the generalizability of the findings to other higher education contexts. Cross-cultural or multi-institutional replication would help examine whether the intention, reflection, and SDL mechanism hold across different learning environments. In addition, the sample exhibited a pronounced gender imbalance, with female students accounting for over 90% of the participants. Prior research in technology acceptance and AI adoption consistently identifies gender as a moderating factor shaping individuals' responses to social influence, trust in AI systems, perceived competence, and affective reactions such as anxiety (Goswami & Dutta, 2016; Møgelvang et al., 2024; Russo et al., 2025).

However, existing studies also emphasize that the direction and strength of these gender-related effects are not uniform, but vary considerably across technologies, levels of technological complexity, learning tasks, and usage contexts (Goswami & Dutta, 2016; Sobieraj & Krämer, 2020). Given the skewed gender composition of the present sample, meaningful gender-based comparisons or moderation analyses could not be conducted. Moreover, the limited gender variability may have attenuated the observable contribution of SN through restricted variance and cohort-specific contextual influences, such as shared peer environments and institutional norms. Consequently, the non-significant effect of SN should be interpreted with caution and should not be generalized to broader higher education populations without further validation using more gender-balanced samples, particularly in light of documented gender-related disparities in AI adoption within higher education contexts (Kalim et al., 2025). Cross-cultural validation could test whether the same intention, reflection, and SDL mechanism holds in diverse learning cultures. Future research could also employ learning analytics or reflective journals to trace how intention and reflection develop over time, providing deeper insights into learners' cognitive regulation. In addition, examining affective factors such as trust in AI, learning satisfaction, or perceived cognitive load (Bonsu & Baffour-Koduah, 2023; Wu & Dong, 2025) could enhance understanding of how emotional and cognitive dimensions jointly shape reflective learning with AI.

6. Conclusion

This study advances the understanding of how first-year university students' intentions to use GenAI are associated with reflective engagement and SDL. By extending the TPB with RT and SDL, the research demonstrates that behavioral engagement acts as a critical mechanism linking motivation to autonomous learning. The ATT and PBC were found to be key predictors of intention, whereas SN showed no significant influence, highlighting the intrinsic and self-regulatory nature of learning based on GenAI. Reflection emerged as a pivotal process associated with meaningful SDL-related experiences in GenAI-supported learning contexts.

The findings contribute theoretically by conceptualizing SDL as a developmental outcome rather than a fixed learner attribute, emphasizing the role of reflection in fostering metacognitive awareness and learner autonomy. Practically, the study recommends lecturers integrating structured reflective activities, such as critique and justification, into instructional design to promote responsible and autonomous GenAI use. Educators and policymakers should embed AI literacy, ethical awareness, and reflective practice into curricula and institutional frameworks to support lifelong learning and academic integrity in using AI.

Despite its strong empirical support, this study acknowledges limitations related to its cross-sectional design and single-university sample. Future research should employ longitudinal and cross-cultural approaches to explore how intention, reflection, and SDL develop over time. Expanding the model to include affective factors such as trust, motivation, and cognitive load would further enrich the understanding of how emotional and cognitive dimensions interact in GenAI-supported learning. Overall, this study underscores that, within a predominantly female student sample, when guided by reflective and self-directed approaches, GenAI can serve as a supportive learning partner in fostering behaviors associated with SDL, thereby laying the foundations of autonomous, ethical, and lifelong learning in university education.

Acknowledgements

The encouragement, collaboration, and support of the VNU University of Education, Vietnam National University, Hanoi, are gratefully acknowledged. This research was supported under grant no. QS.NH.25.27.

AI Statement: Large language models were used for sentence restructuring and grammar correction.

Ethics Statement: Ethical approval was not required for this study, as all participants provided informed consent by proceeding after reading a statement on the study's purpose, anonymity, and the voluntary nature of participation.

Competing Interest Statement: The authors declare that they have no competing interests.

References

Abdallah, N., Katmah, R., Khalaf, K., & Jelinek, H. F. (2025). Systematic review of ChatGPT in higher education: Navigating impact on learning, wellbeing, and collaboration. *Social Sciences & Humanities Open*, *12*, 101866. https://doi.org/10.1016/j.ssaho.2025.101866

Adiguzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology*, *15*(3), ep429. https://doi.org/10.30935/cedtech/13152

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179-211.

Ajzen, I. (2002). Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior. *Journal of Applied Social Psychology*, *32*, 665–683.

Aldammagh, Z., Abdeljawad, R., & Obaid, T. (2021). Predicting mobile banking adoption: An integration of TAM and TBP with trust and perceived risk. *Financial Internet Quarterly*, *17*(3), 35-46. https://doi.org/10.2478/fiqf-2021-0017

Atlas, S. (2023). *ChatGPT for Higher Education and Professional Development: A Guide to Conversational*. https://digitalcommons.uri.edu/cba_facpubs/548

Banh, L., & Strobel, G. (2023). Generative artificial intelligence. *Electronic Markets*, *33*(1), 63. https://doi.org/10.1007/s12525-023-00680-1

Bearman, M., & Ajjawi, R. (2023). Learning to work with the black box: Pedagogy for a world with artificial intelligence. *British Journal of Educational Technology*, *54*(5), 1160-1173. https://doi.org/10.1111/bjet.13337

Bonsu, E. M., & Baffour-Koduah, D. (2023). From the Consumers' Side: Determining Students' Perception and Intention to Use ChatGPT in Ghanaian Higher Education. *Journal of Education, Society & Multiculturalism*, *4*(1), 1-29. https://doi.org/10.2478/jesm-2023-0001

Bordas, A., Le Masson, P., Thomas, M., & Weil, B. (2024). What is generative in generative artificial intelligence? A design-based perspective. *Research in Engineering Design*, *35*(4), 427-443. https://doi.org/10.1007/s00163-024-00441-x

Bosch, C., & Kruger, D. (2024). AI chatbots as Open Educational Resources: Enhancing student agency and Self-Directed Learning. *Italian Journal of Educational Technology*, *32*(1), 53-68.

Boyer, S. L., Edmondson, D. R., Artis, A. B., & Fleming, D. (2013). Self-Directed Learning: A Tool for Lifelong Learning. *Journal of Marketing Education*, *36*(1), 20-32. https://doi.org/10.1177/0273475313494010

Bozkurt, A. (2023). Generative artificial intelligence (AI) powered conversational educational agents: The inevitable paradigm shift. *Asian Journal of Distance Education*, *18*(1). http://www.asianjde.com/ojs/index.php/AsianJDE/article/view/718

Chang, C.-Y., Yang, C.-L., Jen, H.-J., Ogata, H., & Hwang, G.-H. (2024). Facilitating nursing and health education by incorporating ChatGPT into learning designs. *Educational Technology & Society*, *27*(1), 215-230. https://doi.org/10.30191/ETS.202401_27(1).TP02

Chen, L., Chen, P., & Lin, Z. (2020). Artificial Intelligence in Education: A Review. *IEEE Access*, *8*, 75264-75278. https://doi.org/10.1109/ACCESS.2020.2988510

Cheng, E. W. L. (2019). Choosing between the theory of planned behavior (TPB) and the technology acceptance model (TAM). *Educational Technology Research and Development*, *67*(1), 21-37. https://doi.org/10.1007/s11423-018-9598-6

Chiu, T. K. F., Moorhouse, B. L., Chai, C. S., & Ismailov, M. (2024). Teacher support and student motivation to learn with Artificial Intelligence (AI) based chatbot. *Interactive Learning Environments*, *32*(7), 3240-3256. https://doi.org/10.1080/10494820.2023.2172044

Darwin, Rusdin, D., Mukminatien, N., Suryati, N., Laksmi, E. D., & Marzuki. (2024). Critical thinking in the AI era: An exploration of EFL students' perceptions, benefits, and limitations. *Cogent Education*, *11*(1), 2290342. https://doi.org/10.1080/2331186X.2023.2290342

Dehnad, A., Afsharian, F., Hosseini, F., Arabshahi, S. K. S., & Bigdeli, S. (2014). Pursuing a Definition of Self-directed Learning in Literature from 2000–2012. *Procedia - Social and Behavioral Sciences*, *116*, 5184-5187. https://doi.org/10.1016/j.sbspro.2014.01.1097

Esiyok, E., Sahin, G., & Kucukergin, K. G. (2024). Acceptance of Educational Use of AI Chatbots in the Context of Self-Directed Learning with Technology and ICT Self-Efficacy of Undergraduate Students. *International Journal of Human–Computer Interaction*, *41*(1), 641-650. https://doi.org/10.1080/10447318.2024.2303557

Ficko, A., Sarkki, S., Gultekin, Y. S., Egli, A., & Hiedanpää, J. (2025). Reflective thinking meets artificial intelligence: Synthesizing sustainability transition knowledge in left-behind mountain regions. *Geography and Sustainability*, *6*(1), 100257. https://doi.org/10.1016/j.geosus.2024.100257

Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, *18*(1), 39-50. https://doi.org/10.1177/002224378101800104

Franke, G., & Sarstedt, M. (2019). Heuristics versus statistics in discriminant validity testing: a comparison of four procedures. *Internet Research*, *29*(3), 430-447. https://doi.org/10.1108/IntR-12-2017-0515

Garrison, D. R. (1997). Self-directed learning: Toward a comprehensive model. *Adult Education Quarterly*, *48*(1), 18-33. https://doi.org/10.1177/074171369704800103

Gibbons, M. (2002). *The self-directed learning handbook: Challenging adolescent students to excel*. John Wiley & Sons, Amercan.

Gómez-Ramirez, I., Valencia-Arias, A., & Duque, L. (2019). Approach to M-learning Acceptance Among University Students: An Integrated Model of TPB and TAM. *International Review of Research in Open and Distributed Learning*, *20*(3). https://doi.org/10.19173/irrodl.v20i4.4061

Goswami, A., & Dutta, S. (2016). Gender Differences in Technology Usage—A Literature Review. *Open Journal of Business and Management*, *4*, 51-59. https://doi.org/10.4236/ojbm.2016.41006

Hair, J., Black, W., Babin, B., & Anderson, R. (2010). *Multivariate data analysis: A global perspective*. Pearson Education, Upper Saddle River, NJ.

Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., & Ray, S. (2021). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. Third Edition. SAGE Publications.

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. https://doi.org/10.1108/EBR-11-2018-0203

Hendry, G. D., & Ginns, P. (2009). Readiness for self-directed learning: Validation of a new scale with medical students. *Medical Teacher*, *31*(10), 918–920. https://doi.org/10.3109/01421590802520899

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. https://doi.org/10.1007/s11747-014-0403-8

Hiemstra, R. (1988). Self-directed learning: Individualizing instruction. In H. B. L. Associates (Ed.), *Self-directed learning: Application & theory* (pp. 99-124). Athens, GA: Adult Education Department, University of Georgia.

Ho, L.-A., Kuo, T. H., & Lin, B. (2009). Influence of online learning skills in cyberspace. *Internet Research*, *20*(1), 55-71. https://doi.org/10.1108/10662241011020833

Hong, Y.-C., & Choi, I. (2015). Assessing reflective thinking in solving design problems: The development of a questionnaire. *British Journal of Educational Technology*, *46*(4), 848-863. https://doi.org/10.1111/bjet.12181

Kalim, U., Kanwar, A., Sha, J., & Huang, R. (2025). Barriers to AI adoption for women in higher education: a systematic review of the Asian context. *Smart Learning Environments*, *12*(1), 38. https://doi.org/10.1186/s40561-025-00390-5

Kember, D., P., L. D. Y., Alice, J., Yuen, L. A., Jan, M., Kit, S., and Yeung, E. (2000). Development of a Questionnaire to Measure the Level of Reflective Thinking. *Assessment & Evaluation in Higher Education*, *25*(4), 381-395. https://doi.org/10.1080/713611442

Khiat, H. (2017). Academic performance and the practice of self-directed learning: The adult student perspective. *Journal of Further and Higher Education*, *41*(1), 44-59. https://doi.org/10.1080/0309877X.2015.1062849

Knowles, M. (1975). *Self-directed learning: A guide for learners and teachers*. New York: Association Press.

Kock, N. (2017). Common Method Bias: A Full Collinearity Assessment Method for PLS-SEM. In H. Latan & R. Noonan (Eds.), *Partial Least Squares Path Modeling: Basic Concepts, Methodological Issues and Applications* (pp. 245-257). Springer International Publishing. https://doi.org/10.1007/978-3-319-64069-3_11

Kock, N., & Hadaya, P. (2018). Minimum sample size estimation in PLS-SEM: The inverse square root and gamma-exponential methods. *Information Systems Journal*, *28*(1), 227-261. https://doi.org/10.1111/isj.12131

Lee, H.-J. (2005). Understanding and assessing preservice teachers' reflective thinking. *Teaching and Teacher Education*, *21*(6), 699-715. https://doi.org/10.1016/j.tate.2005.05.007

Lee, K., Tsai, P. S., Chai, C. S., & Koh, J. H. L. (2014). Students' perceptions of self-directed learning and collaborative learning with and without technology. *Journal of Computer Assisted Learning*, *30*(5), 425-437. https://doi.org/10.1111/jcal.12055

Li, B., Bonk, C. J., Wang, C., & Kou, X. (2024). Reconceptualizing Self-Directed Learning in the Era of Generative AI: An Exploratory Analysis of Language Learning. *IEEE Transactions on Learning Technologies*, *17*, 1489-1503. https://doi.org/10.1109/TLT.2024.3386098

Li, B., Wang, C., Bonk, C. J., & Kou, X. (2024). Exploring Inventions in Self-Directed Language Learning with Generative AI: Implementations and Perspectives of YouTube Content Creators. *TechTrends*, *68*(4), 803-819. https://doi.org/10.1007/s11528-024-00960-3

Li, Y., Sadiq, G., Qambar, G., & Zheng, P. (2025). The impact of students' use of ChatGPT on their research skills: The mediating effects of autonomous motivation, engagement, and self-directed learning. *Education and Information Technologies*, *30*(4), 4185-4216. https://doi.org/10.1007/s10639-024-12981-9

Lin, X. (2024). Exploring the Role of ChatGPT as a Facilitator for Motivating Self-Directed Learning Among Adult Learners. *Adult Learning*, *35*(3), 156-166. https://doi.org/10.1177/10451595231184928

Møgelvang, A., Bjelland, C., Grassini, S., & Ludvigsen, K. (2024). Gender Differences in the Use of Generative Artificial Intelligence Chatbots in Higher Education: Characteristics and Consequences. *Education Sciences*, *14*(12), 1363. https://doi.org/10.3390/educsci14121363

Monzon, N., & Hays, F. A. (2025). Leveraging Generative Artificial Intelligence to Improve Motivation and Retrieval in Higher Education Learners. *JMIR Med Educ*, *11*, e59210. https://doi.org/10.2196/59210

Ouaazki, A., Bergram, K., Farah, J. C., Gillet, D., & Holzer, A. (2024). *Generative AI-Enabled Conversational Interaction to Support Self-Directed Learning Experiences in Transversal Computational Thinking* Proceedings of the 6th ACM Conference on Conversational User Interfaces, Luxembourg, Luxembourg. https://doi.org/10.1145/3640794.3665542. Article 13, https://doi.org/10.1145/3640794.3665542

Papadakis, S., Lytvynova, S., Mintii, I., Ivanova, S., Selyshcheva, I., & Semerikov, S. (2024). *Advancing lifelong learning with AI-enhanced ICT: A review of 3L-Person 2024.* CEUR Workshop Proceedings.

Parikh, A., Patel, J. D., & Jaiswal, A. K. (2021). Managing job applications online: integrating website informativeness and compatibility in theory of planned behaviour and technology acceptance model. *DECISION*, *48*(1), 97-113. https://doi.org/10.1007/s40622-020-00266-2

Phuong Dung, P. T., An, H. M., Huy, P. Q., & Dinh Quy, N. L. (2023). Understanding the startup's intention of digital marketing's learners: An application of the theory of planned behavior (TPB) and technology acceptance method (TAM). *Cogent Business & Management*, *10*(2), 2219415. https://doi.org/10.1080/23311975.2023.2219415

Rodgers, C. (2002). Defining Reflection: Another Look at John Dewey and Reflective Thinking. *Teachers College Record*, *104*(4), 842-866. https://doi.org/10.1111/1467-9620.00181

Russo, C., Romano, L., Clemente, D., Iacovone, L., Gladwin, T. E., & Panno, A. (2025). Gender differences in artificial intelligence: the role of artificial intelligence anxiety [Original Research]. *Frontiers in Psychology*, *Volume 16 - 2025*. https://doi.org/10.3389/fpsyg.2025.1559457

Saritepeci, M., & Yildiz Durak, H. (2024). Effectiveness of artificial intelligence integration in design-based learning on design thinking mindset, creative and reflective thinking skills: An experimental study. *Education and Information Technologies*, *29*(18), 25175-25209. https://doi.org/10.1007/s10639-024-12829-2

Šedlbauer, J., Činčera, J., Slavík, M., & Hartlová, A. (2024). Students' reflections on their experience with ChatGPT. *Journal of Computer Assisted Learning*, *40*(4), 1526-1534. https://doi.org/10.1111/jcal.12967

Sobieraj, S., & Krämer, N. C. (2020). Similarities and differences between genders in the usage of computer with different levels of technological complexity. *Computers in Human Behavior*, *104*, 106145. https://doi.org/10.1016/j.chb.2019.09.021

Sumuer, E. (2018). Factors related to college students' self-directed learning with technology. *Australasian Journal of Educational Technology*, *34*(4). https://doi.org/10.14742/ajet.3142

Teo, T. (2012). Examining the intention to use technology among pre-service teachers: an integration of the Technology Acceptance Model and Theory of Planned Behavior. *Interactive Learning Environments*, *20*(1), 3-18. https://doi.org/10.1080/10494821003714632

Teo, T., & van Schaik, P. (2012). Understanding the Intention to Use Technology by Preservice Teachers: An Empirical Test of Competing Theoretical Models. *International Journal of Human–Computer Interaction*, *28*(3), 178-188. https://doi.org/10.1080/10447318.2011.581892

Timothy, T., Seng Chee, T., Chwee Beng, L., Ching Sing, C., Joyce Hwee Ling, K., Wen Li, C., & Horn Mun, C. (2010). The self-directed learning with technology scale (SDLTS) for young students: An initial development and validation. *Computers & Education*, *55*(4), 1764-1771. https://doi.org/10.1016/j.compedu.2010.08.001

van den Berg, G., & du Plessis, E. (2023). ChatGPT and Generative AI: Possibilities for Its Contribution to Lesson Planning, Critical Thinking and Openness in Teacher Education. *Education Sciences*, *13*(10). https://doi.org/10.3390/educsci13100998

Venkatesh, V., Morris, M., Davis, G., & Davis, F. (2003). User Acceptance of Information Technology: Toward A Unified View. *MIS Quarterly*, *27*, 425-478. https://doi.org/10.2307/30036540

Venkatesh, V., Thong, J. Y., & Xu, X. (2012). Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS Quarterly*, *36*(1), 157-178.

Wang, C., Li, Z., & Bonk, C. (2024). Understanding self-directed learning in AI-Assisted writing: A mixed methods study of postsecondary learners. *Computers and Education: Artificial Intelligence*, *6*, 100247. https://doi.org/10.1016/j.caeai.2024.100247

Wang, C., Wang, H., Li, Y., Dai, J., Gu, X., & Yu, T. (2024). Factors Influencing University Students' Behavioral Intention to Use Generative Artificial Intelligence: Integrating the Theory of Planned Behavior and AI Literacy. *International Journal of Human–Computer Interaction*, *41*(11), 6649-6671. https://doi.org/10.1080/10447318.2024.2383033

Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. *Expert Systems with Applications*, *252*, 124167. https://doi.org/10.1016/j.eswa.2024.124167

Watson, J. H., & Rockinson-Szapkiw, A. (2021). Predicting preservice teachers' intention to use technology-enabled learning. *Computers & Education*, *168*, 104207. https://doi.org/10.1016/j.compedu.2021.104207

Wei, X., Wang, L., Koszalka, T. A., Lee, L.-K., & Liu, R. (2025). Enhancing pre-service teachers' reflective thinking skills through generative AI-assisted digital storytelling creation: A three-dimensional framework analysis. *Computers & Education*, *235*, 105356. https://doi.org/10.1016/j.compedu.2025.105356

Wu, D., Zhang, S., Ma, Z., Yue, X.-G., & Dong, R. K. (2024). Unlocking Potential: Key Factors Shaping Undergraduate Self-Directed Learning in AI-Enhanced Educational Environments. *Systems*, *12*(9). https://doi.org/10.3390/systems12090332

Wu, H., & Dong, Z. (2025). What Motivates Second Language Majors to Use Generative AI for Informal Learning? Insights From the Theory of Planned Behavior. *IEEE Access*, *13*, 34877-34886. https://doi.org/10.1109/ACCESS.2025.3544489

Yao, Y., Wang, P., Jiang, Y., Li, Q., & Li, Y. (2022). Innovative online learning strategies for the successful construction of student self-awareness during the COVID-19 pandemic: Merging TAM with TPB. *Journal of Innovation & Knowledge*, *7*(4), 100252. https://doi.org/10.1016/j.jik.2022.100252

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

The Impact of AI Literacy on Undergraduate Autonomous Learning

Koravick Thiangtham[1], Ratima Singhachotsukpat[1], Cattareya Wilepana[1], Noawanit Songkram[2] and Jintavee Khlaisang[3]

[1]Department of Educational Technology and Communications, Faculty of Education, Chulalongkorn University, Bangkok, Thailand

[2]Learning Innovation for Thai Society Research Unit (LIFTS), Chulalongkorn University, Bangkok, Thailand

[3]Center of Excellence in Educational Invention and Innovation, Department of Educational Technology and Communications, Faculty of Education, Chulalongkorn University, Bangkok, Thailand

koravick.s@gmail.com
tarnratimas@gmail.com
cattareya.w@gmail.com
noawanit.s@chula.ac.th
jintavee.m@chula.ac.th

https://doi.org/10.34190/ejel.24.2.4188

Abstract: This study explores the factors driving autonomous learning (AU) among undergraduate students in AI-enhanced education. It specifically examines the role of AI literacy (AI-L), critical thinking (CT), self-regulation (SR), and self-efficacy (SE). Data collected from Thai university students were analyzed using Structural Equation Modeling (SEM). The results show that AI-L demonstrated a strong and significant positive influence on all three mediating variables—SE ($\beta = 0.99$, $t = 20.00$), SR ($\beta = 0.93$, $t = 18.53$), and CT ($\beta = 0.70$, $t = 7.30$). SE exerted as the most powerful predictor of AU ($\beta = 0.52$, $t = 6.38$), while critical thinking had a smaller direct impact. The findings suggest that AI-L is a foundational competency that requires metacognitive support. Consequently, educators should utilize strategies like blended learning and reflective practice. These insights encourage a learner-centered approach to digital education, fostering future-ready, autonomous learners.

Keywords: AI literacy, Self efficacy, Self-regulation, Critical thinking, Autonomous learning, Self-directed learning

1. Introduction

Recently, global events and fast-moving technology have transformed education. As a result, there is a renewed focus on how AI literacy (AI-L) shapes a student's ability to learn independently.

While innovations such as blended learning and flipped classrooms have successfully enhanced students' general digital literacy (Latorre-Cosculluela et al., 2021, da Silva Negreiros et al., 2022) A critical area of inquiry is whether AI-L can effectively cultivate autonomous learning (AU)—a competency essential for lifelong learning and aligned with global policy frameworks like the Sustainable Development Goals (SDGs) and the UNESCO Institute for Lifelong Learning (UIL) (2024),While several studies have acknowledged that emerging technologies, including artificial intelligence (AI), can enhance autonomous learning (Intraboonsom, Darasawang and Reinders, 2020, Lai, 2019, Reinders, 2018, Ting, 2015, Papadakis et al., 2024) and foster other positive learning attributes such as self-regulation (Dragomir and Niculescu, 2020, Sinkkonen and Tapani, 2024), self-efficacy (Bewersdorff et al., 2025, Gupta and Jaiswal, 2025, Sun and Shi, 2024, Zhang et al., 2025b), and critical thinking (Lengua-Cantero et al., 2024, Nosratinia and Zaker, 2015, Yüce, 2023), there remains a notable gap regarding whether AI-L directly promotes autonomous learning, or whether its influence is mediated through variables such as self-regulation, self-efficacy, or critical thinking.

Addressing this gap is of significant importance, as autonomous learning enables students to take control of their educational journeys By identifying whether AI-L acts as a standalone skill or a foundational competency supported by mediating factors, this study contributes to the field of educational technology by deepening the theoretical and empirical understanding of AI-L that cultivate self-directed, future-ready learners.

1.1 Research Questions

RQ1. What are the factors influencing AU among undergraduate students?

RQ2. What are the relationships between AI-L, CT, SR, SE and AU among undergraduate students?

Cite this article: Thiangtham, K. et al. 2026. “The Impact of AI Literacy on Undergraduate Autonomous Learning”, *Electronic Journal of e-Learning*, 24(2), pp 32-46, https://doi.org/10.34190/ejel.24.2.4188

RQ3. How do AI-L, CT, SR, and SE influence AU in university students?

RQ4. What is the most effective mediating factor in promoting AU among undergraduate students?

1.2 Research Objectives

This study aims to investigate the factors influencing autonomous learning among undergraduate students by employing SEM. The objectives of this study are as follows:

- To explore the factors influencing AU among undergraduate students
- To examine the relationships between AI-L, CT, SR, SE and AU among undergraduate students.
- To examine the influence of AI-L, CT, SR, and SE to AU in university students
- To identify the most effective mediating factor that promotes AU in undergraduate education.

2. Literature Review and Hypothesis Development

2.1 AI Literacy

AI Literacy (AI-L) refers to the ability to understand and use AI technology effectively. In this context, AI technology includes a diverse suite of technologies ranging from cognitive processing tools (such as Large Language Models- LLM) to specialized pedagogical support system (such as Intelligent Tutoring Systems, AI for image and video production, and Analytical AI). This proficiency encompasses having basic knowledge about AI, evaluating AI performance, assessing AI capabilities, limitations, as well as appropriately creating innovations using AI (Ng et al., 2021b, Ng et al., 2021a, Long and Magerko, 2020, Laupichler et al., 2022, Kandlhofer et al., 2016). In this study, AI-L was categorized into five distinct aspects based on various literature sources.

2.1.1 AI recognition

This includes the ability to distinguish between technological artifacts that use and do not use AI (Long & Magerko, 2020).

2.1.2 AI understanding

This is the ability to explain the concepts and functions of AI and machine learning (Aeri, 2021), along with an understanding of the fundamental principles of AI (Stolpe and Hallström, 2024)

2.1.3 AI leverage

This is the capacity to determine when it is appropriate to utilize AI and when to rely on human judgment and expertise (Long & Magerko, 2020). It also includes the ability to understand and interpret AI-generated outputs with an awareness of potential sources of error or bias (McCoy et al., 2020)

2.1.4 AI ethics

This is the ability to identify key ethical and moral issues related to AI (Long & Magerko, 2020). It also includes the disposition to uphold moral values and ethics in the use and development of AI technologies, along with a willingness to take responsibility for their consequences (Aeri, 2021). Furthermore, it involves an awareness of human-centered considerations (Ng et al., 2021a).

2.1.5 AI utilization

This is the ability to apply AI in addressing emerging and complex problems (Stolpe & Hallström, 2024), and the ability to select and utilize AI tools accurately and appropriately to enhance work efficiency (Chee, Ahn and Lee, 2024).

Educational strategies for AI-L are becoming more diverse, ranging from universal models to discipline-specific integration. Researchers like De Silva et al. (2024) and Kong et al. (2021) advocate for standalone courses designed to establish AI knowledge. De Silva et al. (2024) propose a "Universal AI Literacy Module" developed through human-centered co-design, focusing on four constructs: Foundational Knowledge, Problem Solving, Ethical Practice, and Entrepreneurship, which ensures a baseline of competence. Similarly, Kong et al. (2021) proposed a short, 7-hour standalone course designed to improve AI-L across diverse disciplines without requiring prior programming knowledge. In contrast to standalone models, Southworth et al. (2023) recommend embedding AI learning directly into specific fields of study rather than keeping it separate, such as "GeoAI" for geography, alongside broad certificates. Apart from curriculum structure, Xu et al. (2024) identified that smart classrooms—which incorporate diverse technologies such as AI, virtual reality, digital cameras, and interactive whiteboards to bridge the gap between online and offline learning—can drive successful AI learning. Challenging these competency-based views, Bearman and Ajjawi (2023) argue that instead of trying to teach students to understand how AI works (seeing inside the “black box”), students should learn to work with the "black box" by engaging in meaningful interactions with AI systems, interpreting the technology of AI within the contexts of its use.

2.2 Autonomous Learning

Autonomous Learning (AU) is a concept originating from language education. Started by Henri Holec (1979), who defined the term as "the ability to take charge of one's own learning" (p. 3). In his book titled "Self-instruction in Language Learning," Dickinson (1987) described AU as situations "in which a learner, with others or alone, is working without direct control of a teacher" (p. 5). Later, in 2007, David Little (2007) revisited the fundamentals. He defined autonomy as "the ability to take charge of one's own learning" (p.14), and argued that the concept doesn't have to be specifically applied to language learning in the field.

Early research by Benson and Voller (1997) and Blin (2004) highlighted how technology facilitates self-access and reinforces self-directed practice. Figura and Jarvis (2007) further noted that computer-based media helps learners develop a high-level awareness of their own learning strategies. This concept has expanded in the AI era; Alm (2024) found that Large Language Model tools, such as ChatGPT,allow learners to independently set goals, adjust strategies, and explore language. Similarly, Kalantzis and Cope (2025) proposed the "cyber-social literacy learning" framework, viewing AI as a key driver in creating collaborative human-AI learning spaces that significantly support learner agency and autonomy. However, the mere presence of technology does not automatically guarantee learning autonomy; it requires appropriate instructional design alongside the learner's critical thinking and metacognitive skills. Blin (2004) emphasized that technology only "increases" the opportunity for self-directed practice rather than ensuring it. In the context of AI, Alm (2024) and Kalantzis and Cope (2025) argue that true autonomy depends on the learner's ability to critically evaluate AI-generated content, engage in "intentional decision-making," and utilize effective feedback loops. Furthermore, Thorne (2024) raised concerns regarding academic integrity and inclusivity, suggesting that collaborative teaching strategies are necessary to prevent AI from obscuring the essential mechanisms of the learning process.

Ultimately, despite technological advancements, fostering autonomous learners must remain human-centered. Thorne (2024) asserted the irreplaceable value of human interaction in language learning, warning that AI usage must not diminish the role of the teacher. Meanwhile, Kalantzis and Cope (2025) viewed AI as a transformative force in learning structures but maintained that it should exist to "support, rather than replace, learner agency." The consensus from this research indicates that successful AI integration to promote learner autonomy requires a balanced synthesis of technological capabilities and the critical decision-making of both teachers and learners.

2.3 Self-efficacy

Bandura (1997) defined self-efficacy as the perception of one's capability to organize and execute tasks successfully. Rooted in Social Cognitive Theory, self-efficacy influences how individuals feel, think, and self-motivate. This construct is derived from four key sources: (1) mastery experiences, (2) vicarious experiences, (3) verbal or social persuasion, and (4) physiological and affective states.

Based on this concept, Johnson (2005) and Usher & Pajares (2008) investigated the structural relationships of these four factors with learners. They found that these factors influence behaviors affecting learners' academic success and their choice of major and career. This aligns with Zhou, Chen, and Hou (2022) and Xu, Li, and Yang (2024), who explained that self-efficacy has a positive relationship with self-regulation and motivation. In the context of language learners, those with high self-efficacy tend to use self-regulation strategies effectively, which in turn further stabilizes their self-efficacy. Additionally, Xu et al. (2024) found that self-efficacy plays a mediating role between self-regulated learning and learning engagement. Specifically, learners who self-regulate through communication with instructors and AI, and who perceive they can manage their learning in online environments or smart classrooms, develop confidence in their communication abilities, leading to increased learning engagement (Zhu et al., 2025). Thus, it is evident that self-efficacy has both direct and indirect influences on learning-related variables. As discovered by Honicke & Broadbent (2016) and Sun & Shi (2024), self-efficacy functions as a mediator influencing academic achievement and supporting active learning behaviors, both in higher education contexts (Honicke & Broadbent, 2016) and in offline and online learning environments (Sun & Shi, 2024).

Furthermore, the concept of self-efficacy has expanded into the dimension of confidence in technology use, as explored by Wang & Chuang (2024), Mah & Groß (2024), Chen, Liu & Liu (2024), and Bewersdorff et al. (2025), who studied AI self-efficacy among both learners and instructors. The confidence of learners or instructors in their own ability to understand and use AI effectively serves as a crucial condition linking knowledge, attitudes, and actual usage experience with behavioral outcomes regarding AI use for learning and teaching. Wang and Chuang (2024) examined AI self-efficacy in terms of technological skills, comfort, and human-AI interaction, linking these to motivated learning behaviors. Meanwhile, Bewersdorff et al. (2025) found that AI self-efficacy is related to AI-L, attitudes, and AI usage. In the context of second language learning, Chen and Liu (2024) found that AI self-efficacy positively impacts attitudes, AI usage for language learning, and the reduction of AI anxiety. Additionally, Mah and Groß (2024) found among faculty members that AI self-efficacy correlates with the level of AI adoption in teaching; instructors with positive attitudes and perceived benefits of AI are better able to translate self-efficacy into actual classroom practice.

2.4 Critical Thinking

Critical thinking (CT) has been increasingly recognized as a foundational element in fostering autonomous learning, particularly in higher education contexts. Rezaee and Saleh (2025) describe CT as a purposeful, reflective process shaped by context, culture, and personal experience, aimed at enhancing reasoning, self-awareness, and judgment. Calma and Davies (2025) emphasize CT role in informed decision-making, especially in managerial contexts, while Rivas, Saiz, and Ossa (2022) conceptualize CT as a reasoning-based, problem-solving process geared toward achieving effective outcomes. Other definitions highlight CT as a cognitive skill for logical analysis (Razak et al., 2022), a multifaceted process for understanding and interpreting information (Tathahira, 2020), and a quality of thought that improves communication and engagement with diverse perspectives (Campo et al., 2023, Meneses, Pashchenko and Mikhailova, 2023, Zhang, 2022).

Across literature, CT is consistently linked to the development of Autonomous Learning (AU). Research demonstrates that CT enhances learners' capacity for self-direction, decision-making, and metacognitive reflection. Yüce (2023) identifies CT as a significant predictor of AU ($\beta = .66$, $p < 0.001$), while Nosratinia and Zaker (2015) emphasize how CT allow learners to take responsibility for their learning and manage their progress more effectively. Kravchenko et al. (2023) further support this view by demonstrating that CT development improves communication, analysis, synthesis, and reflection skills, which arecritical components for successful self-study in higher education. Similarly, Lengua-Cantero et al. (2024) highlight how CT fosters self-regulation and deeper understanding, which are essential for navigating the demands of modern education. Horváth (2007) and Iqbal & Akbar (2021) also argue that CT enables learners to plan, assess, and reflect on their goals and learning processes, promoting greater independence and ownership of learning outcomes. Moreover, the integration of critical and creative thinking is increasingly seen as vital, particularly in earlier educational stages. Varías and Callao (2022) propose that a combined model of critical and creative thinking within AU frameworks enhances students' problem-solving abilities and overall educational experience.

Researchers often disagree on whether CT is a general personality trait or a skill tied to specific subjects (Black, 2007). Some scholars argue that CT consists of general skills that can be used in many different areas of life. For example, Dunne (2015) suggests that CT is a way of "being" in the world that combines reasoning with self-reflection and action. Similarly, Bailin (1998) describes it as a process based on universal values like clarity and accuracy, which work across all subject divisions. From this perspective, an ideal critical thinker is someone who has developed a "critical spirit" or a general habit of mind that guides their behavior in all situations. In contrast, other viewpoints suggest that CT is a skill for specific tasks that cannot exist without a particular subject like science or history. One major argument from Black & Dunne (Black, 2007, Dunne, 2015) is that thinking is always "thinking about something," which makes CT conceptually empty if it is not used within a specific field. Even those who see CT as a general ability acknowledge that a person still needs specific knowledge to use these skills well in a particular context (Facione, 1990). For instance, someone may have the mental skills to think critically, but they still need background information to make a good judgment about a complex medical or technical problem. Ultimately, how well a person thinks often depends on their level of experience in a certain area.

2.5 Self-regulation

Self-regulation (SR) is a construct in understanding how individuals manage their cognitive, emotional, and behavioral processes to achieve learning goals. Posner and Rothbart (2000) highlight the complexity of SR, noting its deep connections to volition, genetic predispositions, and social experiences.

While early work often focused on observable behaviors, recent research emphasizes self-regulated learning (SRL) as a dynamic and self-directed process in which learners actively control their educational journey. SRL involves setting goals, selecting strategies, monitoring progress, and reflecting on outcomes—actions driven by internal motivation and belief systems (Blackmore et al., 2021, Higgins, Frankland and Rathner, 2021, Dragomir and Niculescu, 2020, Tadesse et al., 2022, Sinkkonen and Tapani, 2024, Boruchovitch, Simão and Frison, 2023). Zimmerman (2002) outlines this process as a cyclical model comprising forethought, performance, and self-reflection stages, each reinforcing learners' metacognitive awareness and strategic adaptability. Through consistent engagement with SRL processes, learners enhance their autonomy by taking initiative, exercising control over their learning, and making independent decisions. Thus, SR not only improves academic outcomes but also plays a critical role in fostering autonomous learning and preparing students for sustained, lifelong learning.

Based on the literature review, this study proposed the following framework and hypothesis as depicted in Figure 1

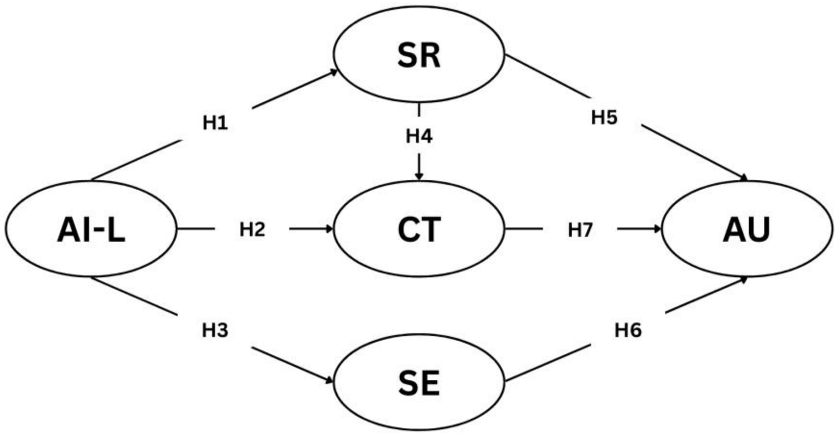


Figure 1: Research model of the study

Hypothesis

The following hypothesis is formulated in this study:

H1: AI positively influences SR, which in turn promotes AU among undergraduate students.

H2: AI positively influences CT, which in turn promotes AU among undergraduate students.

H3: AI positively influences SE, which in turn promotes AU among undergraduate students.

H4: SR positively influences CT, which in turn promotes AU among undergraduate students.

H5: SR positively influences AU among undergraduate students.

H6: SE positively influences CT, which in turn promotes AU among undergraduate students.

H7: CT positively influences AU among undergraduate students.

3. Research Methodology

3.1 Population and Sample Group

The population of this study consists of undergraduate students enrolled in Thai universities. The sample comprises 512 students from diverse academic disciplines—namely Health Sciences, Technology, Engineering, and Social Sciences—across more than 20 universities in Thailand. A convenience sampling method was employed for participant selection. The sample size for conducting SEM was determined based on the recommendation by Hair et al.(2010), resulting in a total of 477 participants, with approximately 9 samples allocated per parameter.

3.2 Research Instrument

The research instrument implemented in this study is a survey questionnaire in an online format. The questionnaire consisted of two parts. Part 1 collected demographic information of the participants, such as age, gender, and department along with their experiences with various types of AI-powered learning tools to understand their background composition. Part 2 of the questionnaire comprised a series of statements designed to examine participants' CT, AI, SR, SE, and their capacity for AU. The participants were asked to rate their agreement with each statement using 5-point Likert scale items ranked from “Strongly agree” (5) to “Strongly disagree” (1). This section of the questionnaire comprised a total of 23 items designed to assess each construct within the research model. Specifically, five items measured AI, five AU, four evaluated CT, four addressed SR, and four measured SE.

3.3 Data Analysis

The proposed research model underwent statistical analysis using SPSS (version 25), and LISREL (8.80) software. The analysis consisted of two stages. In the first stage, Confirmatory Factor Analysis (CFA) was conducted to assess the representativeness of the latent construct with each variable and to assess the fitness of the research model to the actual data to confirm the hypothesized factor structure. The second stage involved conducting a SEM path analysis to examine the causal relationships among the proposed variables. Model fit was evaluated using several indices, including the Comparative Fit Index (CFI), Goodness of Fit Index (GFI), Adjusted Goodness of Fit Index (AGFI), and the Root Mean Square Error of Approximation (RMSEA). According to Hair et al. (2010) an excellent model fit is indicated when the CFI, GFI, and AGFI values exceed 0.95, and the RMSEA value is below 0.05.

4. Results and Findings

4.1 Demographic Profile

The study participants in Figure 2 consisted of 512 respondents, 132 (25.8%) are male, 345 (67.4%) are female and another 35 (6.8%) identified as LGBTQ. Participants came from diverse faculties, with 251 (49.0%) from health sciences, 188 (36.7%) from social sciences, 62 (12.1%) from engineering, and 11 (2.2%) from physical sciences. Additionally, 508 participants (99.2%) reported prior use of Generative AI, while the remaining 4 participants (0.8%) did not.

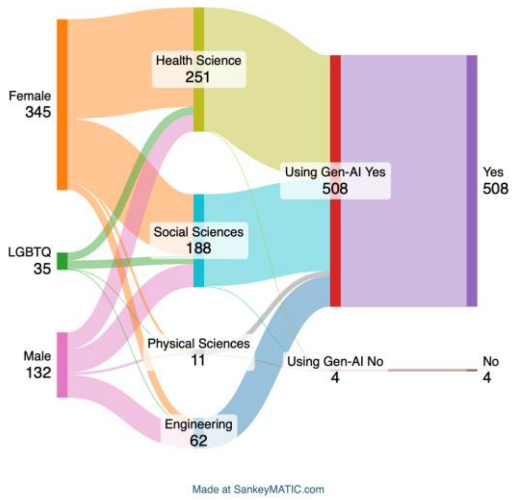


Figure 2: Demographic statistics of the participants

Figure 3 illustrates the usage frequency of generative AI among 508 experienced participants. A total of 150 participants (29.6%) reported using AI 1–3 times per week, followed closely by 145 participants (28.5%) who used it more than six times per week. Additionally, 130 participants (25.6%) used AI 4–6 times per week, 64 participants (12.6%) used it 1–3 times per month, and 19 participants (3.7%) reported using AI less than once per month. Regarding access, most participants—452 (89.0%)—relied on free services, while a smaller portion, 48 participants (9.4%), paid for access either through per-use fees or subscriptions. The remaining 8 participants (1.6%) received access through organizational support.

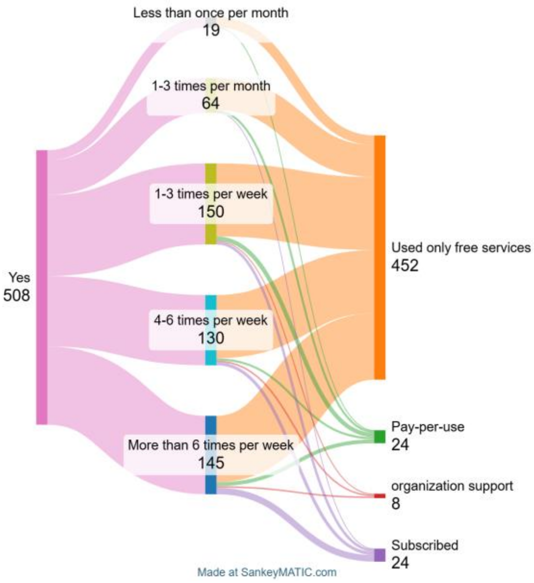


Figure 3: AI usage frequency and type of service

4.2 Measurement Model

In evaluating the measurement model, all constructs demonstrated acceptable to excellent levels of model fit. The SR construct showed a good fit, with a Chi-Square value of 4.84 (df = 2), RMSEA of 0.053, SRMR of 0.008, and fit indices including CFI and GFI at 1.00, and AGFI at 0.98. Similarly, SE construct exhibited strong model fit, reflected by a Chi-Square of 4.00 (df = 3), RMSEA of 0.026, SRMR of 0.005, and perfect values for CFI, GFI (1.00), and a high AGFI of 0.98. AI construct also indicated a good fit with a Chi-Square of 5.58 (df = 2), RMSEA of 0.05, SRMR of 0.01, and high indices (CFI = 0.99, GFI = 0.99, AGFI = 0.97). CT showed an excellent fit with a Chi-Square of 0.23 (df = 1), RMSEA of 0.00, SRMR of 0.002, and perfect CFI and GFI values (1.00), along with AGFI of 0.99. Lastly, AU displayed strong model fit indicators, including a Chi-Square of 1.18 (df = 2), RMSEA of 0.00, SRMR of 0.004, and perfect CFI and GFI scores (1.00), with AGFI at 0.99.

Table 1 presents psychometric properties of five latent constructs measured by multiple items, including their means (M), standard deviations (SD), skewness (SK), kurtosis (KU), factor loadings (FL), and key validity and reliability indices (Cronbach's alpha, Construct Reliability, and Average Variance Extracted). In summary, the measurement model shows strong psychometric validity across all constructs. AI, CT, SE, SR and AU. Each demonstrates high factor loadings, excellent construct reliability (CR > 0.70), and solid average variance extracted (AVE > 0.50). Although AI-L has slightly lower factor loadings and AVE (0.597), it still meets acceptable standards, indicating sound measurement quality overall. Correlation analyses were also applied within the scope of studying the connection between AI, CT, SR, SE and AU, as depicted in Table 2:

Table 1: Descriptive statistics of the measurement model

Construct	Items	Questions	M	SD	SK	KU	FL	α	CR	AVE
AU	AU1	I can set learning goals on a daily, weekly, or long-term basis to support my own learning.	3.66	0.94	-.455	.092	0.82	0.931	0.744	0.935
	AU2	I plan methods and manage my time effectively to achieve learning success.	3.75	0.90	-.400	-.075	0.87			

Construct	Items	Questions	M	SD	SK	KU	FL	α	CR	AVE
	AU3	I am capable of self-regulated learning without relying on others.	3.81	0.90	-.450	-.023	0.86			
	AU4	I can select learning styles, techniques, and technologies that suit my personal needs.	3.90	0.90	-.521	-.061	0.87			
	AU5	I am able to self-assess in order to improve the effectiveness of my learning.	3.84	0.89	-.390	-.196	0.89			
CT	CT1	I am able to analyze information in detail to understand and draw conclusions.	3.79	0.87	-.450	.226	0.81	0.925	0.902	0.698
	CT2	I can use reasoning to compare opinions or experiences to support well-informed decisions.	3.86	0.88	-.552	.276	0.86			
	CT3	I am capable of establishing criteria to assess importance or value.	3.80	0.87	-.499	.332	0.84			
	CT4	I can apply reasoning and skills to find solutions to various problems.	3.89	0.89	-.527	.127	0.83			
SR	SR1	I can select and apply strategies to enhance effective learning.	3.83	0.89	-.586	.539	0.80	0.923	0.913	0.723
	SR2	I am intrinsically motivated to set goals, engage in learning, persevere through challenges, and	3.83	0.87	-.412	.065	0.86			

Construct	Items	Questions	M	SD	SK	KU	FL	α	CR	AVE
		believe in my own abilities.								
	SR3	I can adapt and improve my learning strategies based on feedback and self-reflection.	3.81	0.90	-.414	-.027	0.86			
	SR4	I actively participate in collaborative learning with others in the classroom.	3.88	0.95	-.675	.355	0.88			
SE	SE1	I have sufficient knowledge to independently solve complex problems in learning or work.	3.74	0.85	-.324	.100	0.77	0.938	0.926	0.715
	SE2	I believe that my existing skills enable me to succeed in both simple and challenging tasks.	3.83	0.87	-.441	.075	0.89			
	SE3	I can plan and manage my work effectively to achieve set goals.	3.87	0.89	-.478	.010	0.88			
	SE4	Even when facing obstacles, I remain confident in my ability to find solutions.	3.88	0.88	-.513	.167	0.88			
	SE5	I am confident in my capacity to develop diverse knowledge and skills to meet future challenges	3.88	0.87	-.455	.020	0.80			
AI	AI1	I can identify which technologies utilize AI.	3.50	1.00	-.247	-.304	0.73	0.895	0.880	0.597
	AI2	I am able to explain the basic principles of how AI works.	3.280	1.00	.000	-.460	0.67			

Construct	Items	Questions	M	SD	SK	KU	FL	α	CR	AVE
	AI3	I can analyze the advantages and limitations of AI.	3.69	0.92	-.433	.106	0.83			
	AI4	I am aware of ethical issues (e.g., bias, privacy) and potential societal impacts arising from AI use.	3.86	0.96	-.543	-.094	0.82			
	AI5	I can apply AI tools (e.g., ChatGPT, AI image generators) appropriately to different types of tasks.	3.94	0.96	-.758	.330	0.80			

Table 2: Inter-construct correlation

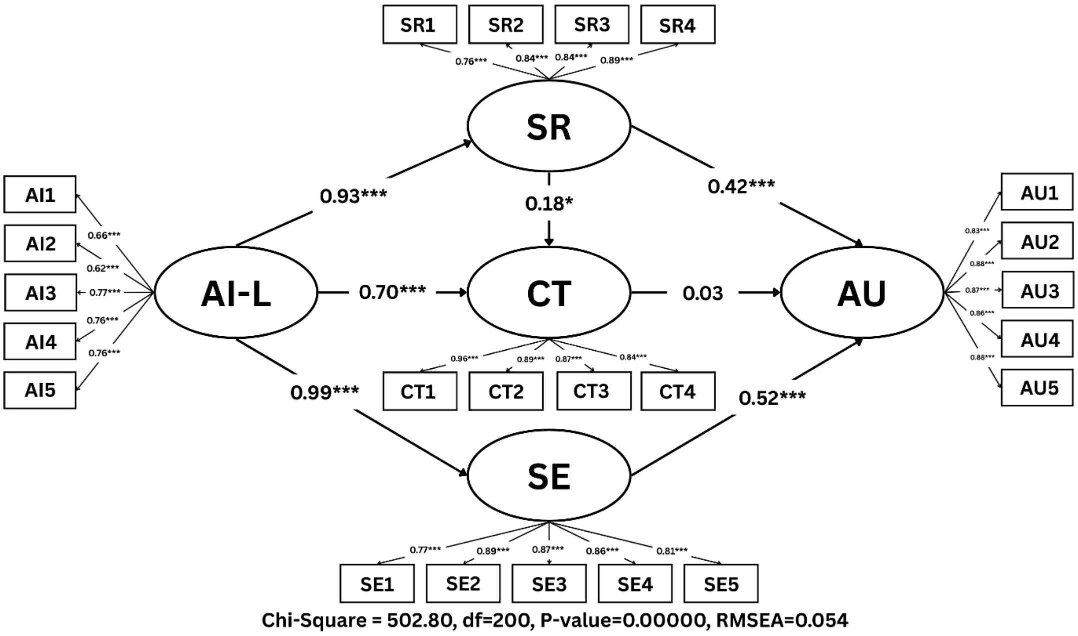
	AI-L	SR	SE	CT	AU
AI-L	1				
SR	.494	1			
SE	.549	.055	1		
CT	.213	.708	.575	1	
AU	.768	.490	.885*	.698	1

**Correlation is significant at the 0.05 level (2-tailed)*

4.3 Structural Model: Goodness of fit Statistics and Hypothesis Testing

As depicted in Figure 4, The SEM analysis revealed that the model fits the data well, with strong indices such as CFI (0.99), RMSEA (0.054), and SRMR (0.026). Although GFI (0.92) and AGFI (0.89) were slightly below ideal thresholds, they remained within acceptable ranges. A key finding was the significant influence of AI-L on all three mediators—SE (β = 0.99), SR (β = 0.93), and Critical Thinking (β = 0.70)—indicating that greater AI proficiency is closely linked to stronger self-directed learning capacities and analytical skills. Among predictors of AU, SE had the strongest direct effect (β = 0.52) whereas SR and CT did not show significant contributions to the development of AU.

The SEM analysis confirms hypotheses H1 through H6, demonstrating that AI has a significant influence on SE, SR and CT. Additionally, significant paths were observed from SE to AU and from SR to AU. In contrast, hypothesis H7 is not supported, indicating no significant relationship between CT and AU, as illustrated in Table 3.



**p < 0.05, ** p < 0.01. *** p < 0.001.*

Figure 4: Results of Structural Equation Model

Table 3: Structural Equation Modelling results of the proposed model

Hypothesis	Standardized Solution	Results
H1: AI-L → SR	0.93***	Supported
H2: AI-L → CT	0.70***	Supported
H3: AI-L → SE	0.99***	Supported
H4: SR → CT	0.18*	Supported
H5: SR → AU	0.42***	Supported
H6: SE → AU	0.52***	Supported
H7: CT → AU	0.03	Not supported

**p < 0.05, ** p < 0.01. *** p < 0.001.*

5. Discussion and Implications

One of the most significant findings of this study is the strong impact of AI-L on SE, which in turn significantly predicts AU. Students who develop a foundational knowledge of AI-L concepts and tools gain confidence in using AI effectively, reducing anxiety and enhancing their belief in their ability to succeed. This supports Bandura's (1997) theory that that succeeding at tasks strengthens self-efficacy, which later confirmed by Chen et al. (2024) and Zhang et al. (2025). The strong impact of AI-L on SE also reinforces the importance of equipping students with essential AI knowledge to boost self-confidence and learning autonomy (Bewersdorff et al., 2025; Gupta & Jaiswal, 2025). To translate these findings into practice, educators should add basic AI lessons to all subjects, using flexible learning plans that fit each major. The use of AI-enabled assessment tools, digital credentials, and faculty development programs, including hands-on workshops and case-based training, can further support effective implementation, as suggested by Mah and Groß (2024).

The study also highlights the role of AI-L in fostering SR, which contributes to the development of AU. As students engage with AI tools for the purpose of education such as intelligent tutoring systems and generative AI—they are prompted to make decisions, monitor progress, and assess outputs, all of which reinforce SR. This aligns with research on flipped and blended learning (Yoon, Hill and Kim, 2021, Latorre-Cosculluela et al., 2021), which supports student autonomy and metacognitive skill development. Educators are encouraged to adopt models like the MOOC-based flipped classroom (Pérez-Sanagustín et al., 2021) or the FCM-AI integration (Zhu et al., 2025) to promote deeper learning and critical reflection. Additionally, the proceedings from the 3L-Person 2024 workshop (Papadakis et al., 2024) support this connection. The Person-oriented Approach central to the workshop highlights how AI-enhanced ICT tools can support individualized learning paths. Specifically, research within this volume demonstrates that such environments can foster a culture of independence through formative assessment and promote successful professional self-realization by developing digital competencies in immersive environments.

A noteworthy observation from our study is that AI usage does not strongly CT, nor does CT significantly influence AU. The relationship between CT and AI-L is complex and context-dependent, leading to mixed results depending on whether CT is seen as a result of using AI or a requirement for acquiring AI competency. Gerlich (2025) and Zhou et al. (2024) found that high trust in AI, particularly among younger users or consumers of paid AI services, often leads to cognitive offloading and hinders critical reflection. In contrast, Hornberger et al. (2023) frame CT as a core component of AI competence. This has been empirically confirmed via Confirmatory Factor Analysis ($X^2/df = 2.54$) that critical interpretation and technical knowledge are driven by the same latent literacy capability. In summary, there is a strong connection if CT is seen as a skill needed to understand AI. However, if it is seen as a result of using AI, the connection is inconsistent and depends on other factors like the user's age and the type of service.Another point worth discussing is the weak relationship between CT and AU which interpret that having higher-order thinking does not guarantee autonomous learning. This finding contrasts with previous research, where scholars have agreed that CT and AU share a strong relationship (Campo et al., 2023, Kravchenko et al., 2023, Yüce, 2023). Typically, learners with higher CT tend to be more self-directed. Similarly, it is well-established that autonomous learners often possess better higher-order thinking skills than those who rely on others for their learning. We suggest that the context of using AI as a "learning companion" explains this discrepancy. Even if students develop CT (mainly to critique AI's output) through this process, this might not translate into self-directed learning because they may perceive using AI as relying on an external source. Therefore, even learners with high CT might assess themselves as "reliant on AI" when they use these tools. Recent research highlights the negative effects of relying too heavily on AI. Zhang et al. (2025a) found that excessive dependence causes a decline in independent thinking skills, which leads to reduced autonomy. Similarly, Aljuraid (2025) noted that heavy reliance affects professionals, such as clinicians, by diminishing their sense of control which can lead to a loss of identity. Furthermore, Tian and Zhang (2025) explain that relying on AI promotes cognitive offloading which weakens their ability to think independently over time. These findings are important for teachers who want to apply this research in the classroom. It suggests that teachers, in cooperation with their institutions and faculties, should set clear boundaries for AI usage. By defining what level of use is acceptable, educators can help students feel more self-directed. This approach encourages students to become confident and responsible learners, without feeling that they are relying too heavily on AI.

It must be acknowledged that a debate is still being held by researchers about whether CT should be conceptualized as a domain-specific skill or as a transferable attribute applicable across diverse disciplines. For the purposes of this study, CT was scoped to a generalized predisposition shaped by a learner's formative upbringing and learning experiences. Because of this, specific types of CT, like the logic that is needed to write AI prompts or the skills that are learned from using AI tools over time, were not tested. It is therefore suggested that the link between CT and AU might be different if it were studied as a skill for specific tasks, rather than the general habit that was used for this work.

6. Conclusion

This study aimed to explore the factors influencing CT among undergraduate students, with a particular focus on the roles of AI-L, CT, SR, and SE. Through structural equation modeling (SEM). AI-L was found to have a significant and robust positive effect on all three mediating variables—SE ($\beta = 0.99$, $t = 20.00$), SR ($\beta = 0.93$, $t = 18.53$), and CT ($\beta = 0.70$, $t = 7.30$). Among the predictors of AU, SE emerged as the most influential factor ($\beta = 0.52$, $t = 6.38$), highlighting how important it is for academic achievement. These findings offer valuable implications for promoting AI-L, particularly through pedagogical models such as blended learning and flipped classrooms, both of which are well-documented for enhancing self-regulatory behaviors and learner confidence.

While critical thinking showed a relatively smaller direct effect on academic outcomes, its presence remains crucial in shaping comprehensive AI-L. It is important to design classes where students learn to check AI outputs for errors instead of relying on them completely. By linking AI-L to metacognitive skills, this study broadens the scope of e-learning research. It suggests that AI-powered learning tools must be paired with reflective instruction and ongoing feedback. This approach helps educators design student-centered curricula that foster true autonomy and digital readiness.

7. Limitation and Suggestion for Future Studies

This study contains limitations that may influence the interpretation of the data. Therefore, researchers should approach our findings with caution. The first limitation concerns the study context. We conducted this research among Thai undergraduate students within a specific educational and cultural setting. The second limitation involves reliance on self-report methods. This approach might introduce subjective bias, as participants may inaccurately assess their own skills. Consequently, generalizing these findings to other populations requires careful consideration.

To build upon the findings of this research, we recommend that future studies further explore the relationship between critical thinking, AI usage, and autonomous learning. As noted, evidence regarding this relationship remains contradictory across the literature. Additionally, there is a growing need for research focused on the development of curricula and instructional designs within specific subject areas. These pedagogical interventions should aim to enhance self-efficacy through the cultivation of AI-L, meeting the demands of 21st-century education in an AI-driven world.

Acknowledgement

We would like to express our deepest gratitude to Professor Dr. Noawanit Songkram and Professor Dr. Jintavee Khlaisang for their continued support, invaluable guidance, and encouragement throughout the development of this paper. Their mentorship and expertise have been instrumental in shaping the direction and quality of our research.

Declaration of AI use: This manuscript involved the use of AI Writing Assistants (Grammarly) and Generative AI (Google Gemini) solely for the purpose of language editing and grammar refinement. The content, ideas, structure, and interpretations presented in the article are entirely the author's original work. AI assistance was limited to improving clarity, coherence, and academic style without altering the substantive meaning of the text.

Ethics Statement: This study strictly adheres to the ethical guidelines established by Chulalongkorn University, Thailand. All respondents recruited for the questionnaire were adults aged 18 years or older and participated on a purely voluntary basis. Prior to data collection, participants were fully briefed on the research objectives and the intended use of the findings. Informed consent was obtained from all respondents before they accessed the questionnaire, confirming their understanding of the study's nature. To protect participant privacy, all responses have been anonymized, and respondents were informed of their right to withdraw from the study at any stage. Furthermore, the researcher declares that there are no known conflicts of interest associated with any organization.

References

Aeri, L. (2021) "The Effect of ARTIFICIAL INTELLIGENCE Literacy Education on University Students Ethical Consciousness of Artificial Intelligence," *Robotics & AI Ethics*, 6(3), pp. 52–61. Available at: https://doi.org/10.22471/ai.2021.6.3.52.

Aljuraid, R. (2025) "The Illusion of Control: AI Chatbot Dependency and the Threat to Clinical Autonomy," *Studies in health technology and informatics*, 332, pp. 211–215. Available at: https://doi.org/10.3233/SHTI251529.

Bailin, S. (1998) Skills, Generalizability and Critical Thinking, Paideia Archive: 20th World Congress of Philosophy.

Bandura, A. and Wessels, S. (1997) *Self-efficacy*. Cambridge University Press Cambridge.

Bewersdorff, A. et al. (2025) "AI advocates and cautious critics: How AI attitudes, AI interest, use of AI, and AI literacy build university students' AI self-efficacy," *Computers and Education: Artificial Intelligence*, 8, p. 100340. Available at: https://doi.org/10.1016/j.caeai.2024.100340.

Black, B. (2007) 'Critical Thinking – a tangible construct?', *Research Matters: A Cambridge Assessment Publication*, 3, pp. 2–4. Available at: http://www.cambridgeassessment.org.uk/research-matters/

Blackmore, C. et al. (2021) "A Review of Self-Regulated Learning and Self-Efficacy: The Key to Tertiary Transition in Science, Technology, Engineering and Mathematics (STEM)," *International Journal of Higher Education*, 10(3), pp. 169–177.

Boruchovitch, E., Simão, A.M.V. and Frison, L.M.B. (2023) "Contributions, limitations, and challenges of self-regulated learning in the educational context," *ETD: Educaçao Temática Digital*, (25), p. 34.

Calma, A. and Davies, M. (2025) "Assessing students' critical thinking abilities via a systematic evaluation of essays," *Studies in Higher Education*, pp. 1–16.

Campo, L. et al. (2023) "Methodologies for fostering critical thinking skills from university students' points of view," *Education Sciences*, 13(2), p. 132.

Chee, H., Ahn, S. and Lee, J. (2024) "A Competency Framework for AI Literacy: Variations by Different Learner Groups and an Implied Learning Pathway," *British Journal of Educational Technology* [Preprint]. Available at: https://doi.org/10.1111/bjet.13556.

Dickinson, L. (1987) *Self-instruction in language learning*. Cambridge University Press (System).

Dragomir, I. and Niculescu, B.-O. (2020) "Different Approaches to Developing Writing Skills," *Land Forces Academy Review*, 25, pp. 201–206. Available at: https://doi.org/10.2478/raft-2020-0024.

Dunne, G. (2015) "Beyond critical thinking to critical being: Criticality in higher education and life," *International Journal of Educational Research*, 71, pp. 86–99.

Facione, P.A. (1990) "Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction. Research findings and recommendations.". Available at: https://eric.ed.gov/?id=ED315423

Gupta, S. and Jaiswal, R. (2025) "A deep learning-based hybrid PLS-SEM-ANN approach for predicting factors improving AI-driven decision-making proficiency for future leaders," *Journal of International Education in Business*, 18(2), pp. 234–268. Available at: https://doi.org/10.1108/jieb-05-2024-0058.

Hair, J.F., Black, W.C., Babin, B.J. and Anderson, R.E. (2010) *Multivariate Data Analysis: A Global Perspective*. 7th ed. Upper Saddle River, NJ: Pearson Education.

Higgins, N., Frankland, S. and Rathner, J. (2021) "Self-regulated learning in undergraduate science," *International Journal of Innovation in Science and Mathematics Education*, 29(1).

Holec, H. (1981*) Autonomy and Foreign Language Learning*. Oxford: Pergamon Press Horváth, I. (2007) "Autonomous learning: what makes it work in postgraduate interpreter training?," *Across Languages and Cultures*, 8(1), pp. 103–122.

Intraboonsom, C., Darasawang, P. and Reinders, H. (2020) "Teacher's Practices in Fostering Learner Autonomy in a Thai University Context," *Journal of Language Teaching and Research*, 11(2), pp. 194–203.

Iqbal, N. and Akbar, R.A. (2021) "Effect of Autonomous Learner Approach on Prospective Teachers' Critical Thinking Skills," *Global Educational Studies Review*, VI, 6, pp. 369–375.

Kandlhofer, M. et al. (2016) "Artificial intelligence and computer science in education: From kindergarten to university," in. 2016 IEEE Frontiers in Education Conference (FIE), pp. 1–9. Available at: https://doi.org/10.1109/FIE.2016.7757570.

Kravchenko, O. et al. (2023) "The use of technology-based model of critical thinking development to reshape students' self-study process," *European Journal of Educational Research*, 1(12), pp. 281–296.

Lai, C. (2019) "Technology and Learner Autonomy: An Argument in Favor of the Nexus of Formal and Informal Language Learning," *Annual Review of Applied Linguistics*, 39, pp. 52–58. Available at: https://doi.org/10.1017/S0267190519000035.

Laupichler, M.C. et al. (2022) "Artificial intelligence literacy in higher and adult education: A scoping literature review," *Computers and Education: Artificial Intelligence*, 3, p. 100101. Available at: https://doi.org/10.1016/j.caeai.2022.100101.

Little, D. (2007) "Language Learner Autonomy: Some Fundamental Considerations Revisited," *Innovation in Language Learning and Teaching*, 1(1), pp. 14–29. Available at: https://doi.org/10.2167/illt040.0.

Long, D. and Magerko, B. (2020) "What is AI Literacy? Competencies and Design Considerations," in. *ACM.* Available at: https://doi.org/10.1145/3313831.3376727.

McCoy, L.G. et al. (2020) "What do medical students actually need to know about artificial intelligence?," *Digital Medicine*, 3(1). Available at: https://doi.org/10.1038/s41746-020-0294-7.

Meneses, L.F.S., Pashchenko, T. and Mikhailova, A. (2023) "Critical thinking in the context of adult learning through PBL and e-learning: a course framework. Thinking Skills Creativity 1, 101358 (2023)."

Ng, D.T.K., Leung, J.K.L., Chu, K.W.S., et al. (2021) "AI Literacy: Definition, Teaching, Evaluation and Ethical Issues," Proceedings of the Association for Information Science and Technology, 58(1), pp. 504–509. Available at: https://doi.org/10.1002/pra2.487.

Ng, D.T.K., Leung, J.K.L., Chu, S.K.W., et al. (2021) "Conceptualizing AI literacy: An exploratory review," *Computers and Education: Artificial Intelligence*, 2, p. 100041. Available at: https://doi.org/10.1016/j.caeai.2021.100041.

Papadakis, S. et al. (2024) "Advancing lifelong learning with AI-enhanced ICT: A review of 3L-Person 2024," in. CEUR Workshop Proceedings.

Pawar, M. and Rai, R.D. (2023) "Ethical Challenges of Artificial Intelligence in Education: Achieving Learner Centricity with Respect to Learner Autonomy," International Conference on Computers in Education [Preprint]. Available at: https://doi.org/10.58459/icce.2023.1437.

Pellas, N. (2025) "The role of students' higher-order thinking skills in the relationship between academic achievements and machine learning using generative AI chatbots," *Research and Practice in Technology Enhanced Learning*, 20. Available at: https://doi.org/10.58459/rptel.2025.20036.

Pérez-Sanagustín, M. et al. (2021) "A MOOC-based flipped experience: Scaffolding SRL strategies improves learners' time management and engagement," *Computer Applications in Engineering Education*, 29(4), pp. 750–768.

Posner, M.I. and Rothbart, M.K. (2000) "Developing mechanisms of self-regulation," *Development and psychopathology*, 12(3), pp. 427–441.

Razak, A.A. et al. (2022) “Improving critical thinking skills in teaching through problem-based learning for students: A scoping review,” *International Journal of Learning, Teaching and Educational Research*, 21(2), pp. 342–362.

Reinders, H. (2018) “Technology and autonomy,” *The TESOL encyclopedia of English language teaching*, 15(3), pp. 1–5.

Rezaee, A.A. and Saleh, E. (2025) “Critical thinking in the cultural components of EFL textbooks: CCR model and its application,” *Thinking Skills and Creativity*, 57, p. 101782. Available at: https://doi.org/10.1016/j.tsc.2025.101782.

Rivas, S.F., Saiz, C. and Ossa, C. (2022) “Metacognitive strategies and development of critical thinking in higher education,” *Frontiers in Psychology*, 13, p. 913219.

da Silva Negreiros, F.D. et al. (2022) “E-MunDiabetes: A Mobile Application for Nursing Students on Diabetes Education During the COVID-19 Pandemic,” *CIN - Computers Informatics Nursing*, 40(5), pp. 325–334. Available at: https://doi.org/10.1097/CIN.0000000000000881.

Sinkkonen, M. and Tapani, A. (2024). “Review of the Concept ‘Self-Regulated Learning’: Defined and Used in Different Educational Contexts,” *International journal on social and education sciences*, 6(1), pp.130–151. Available at: https://doi.org/10.46328/ijonses.640.

Stolpe, K. and Hallström, J. (2024) “Artificial intelligence literacy for technology education,” *Computers and Education Open*, 6, p. 100159. Available at: https://doi.org/10.1016/j.caeo.2024.100159.

Sun, W. and Shi, H. (2024) “Fostering success in online English education: Exploring the effects of ICT literacy, online learning self-efficacy, and motivation on deep learning,” *Education and Information Technologies*, 29(18), pp. 24899–24920. Available at: https://doi.org/10.1007/s10639-024-12827-4.

Tathahira, T. (2020) “Promoting students’ critical thinking through online learning in higher education: Challenges and strategies,” *Englisia: Journal of Language, Education, and Humanities*, 8(1), pp. 79–92.

Tian, J. and Zhang, R. (2025) “Learners’ AI dependence and critical thinking: The psychological mechanism of fatigue and the social buffering role of AI literacy,” *Acta Psychologica*, 260. Available at: https://doi.org/10.1016/j.actpsy.2025.105725.

Ting, Y.-L. (2015) “Tapping into students’ digital literacy and designing negotiated learning to promote learner autonomy,” *The Internet and Higher Education*, 26, pp. 25–32. Available at: https://doi.org/10.1016/j.iheduc.2015.04.004.

United Nations (2024) Goal 4: Quality Education. Available at: https://globalgoals.org/goals/4-quality-education/

Varías, I. and Callao, M. (2022) “Autonomous learning strategies: critical and creative thinking in primary education.”

Xu, J., Li, J. and Yang, J. (2024) “Self-regulated learning strategies, self-efficacy, and learning engagement of EFL students in smart classrooms: A structural equation modeling analysis,” *System*, 125, p. 103451. Available at: https://doi.org/10.1016/j.system.2024.103451.

Yang, C., Wei, M. and Liu, Q. (2025) “Intersections between cognitive-emotion regulation, critical thinking and academic resilience with academic motivation and autonomy in EFL learners: Contributions of AI-mediated learning environments,” *British Educational Research Journal* [Preprint]. Available at: https://doi.org/10.1002/berj.4140.

Yoon, M., Hill, J. and Kim, D. (2021) “Designing supports for promoting self-regulated learning in the flipped classroom,” *Journal of Computing in Higher Education*, 33, pp. 398–418.

Yüce, E. (2023) “Critical thinking, autonomous learning, and academic grit among preservice EFL teachers,” Thinking Skills and Creativity, 50, p. 101382. Available at: https://doi.org/10.1016/j.tsc.2023.101382.

Zhang, D. et al. (2025) “Exploring the relationship between AI literacy, AI trust, AI dependency, and 21st century skills in preservice mathematics teachers,” *Scientific Reports*, 15(1). Available at: https://doi.org/10.1038/s41598-025-99127-0.

Zhang, Dongli et al. (2025) “Exploring the relationship between AI literacy, AI trust, AI dependency, and 21st century skills in preservice mathematics teachers,” *Scientific Reports*, 15(1), p. 14281. Available at: https://doi.org/10.1038/s41598-025-99127-0.

Zhang, Q. et al. (2025) “Exploring the Dynamics of Artificial Intelligence Literacy on English as a Foreign Language Learners’ Willingness to Communicate: The Critical Mediating Roles of Artificial Intelligence Learning Self-Efficacy and Classroom Anxiety,” *Behavioral Sciences*, 15(4), p. 523. Available at: https://doi.org/10.3390/bs15040523.

Zhang, Y. (2022) “The research on critical thinking teaching strategies in college English classroom,” *Creative Education*, 13(4), pp. 1469–1485.

Zhu, Q. et al. (2025) “Flipped classroom for music theory course in Chinese universities: a conceptual framework of practice,” *Cogent Education*, 12(1), p. 2498146.

Zimmerman, B.J. (2002) “Becoming a self-regulated learner: An overview,” *Theory into practice*, 41(2), pp. 64–70.

Ecological Predictors of AI Literacy in Chinese K-12 Teachers: A Structural Equation Modeling Study

Xiaofan Wu[1] and Nagaletchimee Annamalai[2,3]
[1]School of Distance Education, Universiti Sains Malaysia, Malaysia
[2]Department of Languages and Culture, College of Humanities and Sciences, Ajman University, UAE
[3]Research Center (HSSRC), Ajman University, UAE

wuxiaofan_0113@163.com
a.annamalai@ajman.ac.ae (corresponding author)

https://doi.org/10.34190/ejel.24.2.4155

Abstract: Although AI is being rapidly developed and applied in education, gaps remain in factors affect teachers' AI literacy. A cross-sectional survey of 1,680 teachers was conducted to explore relationships between school environment, social environment, teacher self-efficacy, and AI literacy via structural equation modeling (CFI = 0.986; RMSEA = 0.03). The results showed that teachers' AI literacy was 3.89 ± 1 (out of 5) in total, and the theory-practice gap was significant: stronger performance in awareness (β = 0.75) and ethics (β = 0.76), but weaker performance in application literacy (β = 0.72) and evaluation literacy (β = 0.81). School environment had the strongest direct effect on AI literacy (β = 0.270, p < 0.001), followed by teacher self-efficacy, which served as an important mediator (β = 0.259, p < 0.001). Social environment had no direct effect on teachers' AI literacy (β = 0.060, p = 0.362), implying that distal effects need to be mediated by school. Demographic analysis showed urban–rural differences, decline after age 40, and subject differences (science > liberal arts). Therefore, we suggest that policymakers should transfer to supporting school-level interventions with targeted resources allocation. School leaders should create supportive technological environments and self-efficacy programs. In addition, teachers should participate in hands-on training with a focus on practical skills. This study provides useful references for integrating AI into K-12 education in China.

Keywords: AI literacy, AI for teacher education, Teacher self-efficacy, Ecological systems theory

Highlights

1. Social environment influences teachers' AI literacy indirectly through self-efficacy's mediating role, demonstrating interplay between contextual factors and personal beliefs.
2. Teacher's self-efficacy significantly mediates between social environment and AI literacy, indicating how confidence in capabilities facilitates technology adoption.
3. K-12 teachers exhibit notable deficiencies in AI application and evaluation competencies, revealing challenges in practical implementation and critical assessment.
4. School teaching environment directly impacts teachers' AI literacy, highlighting the importance of institutional support, resources, and professional development opportunities.

1. Introduction

Artificial intelligence (AI) is revolutionizing education worldwide by providing tools to personalize learning and enhance teacher efficiency (Su & Ng, 2022). UNESCO has called AI literacy essential for sustainable educational ecosystems and recommended that countries prepare educators for AI teaching (Pedro et al., 2019).

However, teacher AI preparedness is remarkably different across systems. For instance, while in Nordic countries such as Finland, an integrated AI education framework has been developed (and in Singapore a nationwide teacher AI literacy has been put into practice) (Chai et al., 2021), other systems are far behind. For example, only 23% of teachers in the US felt prepared to teach with AI even though 60% indicated interest (Education Week Research Center, 2023). For European systems, research indicated great variation in digital competency framework development, with Estonia and Netherlands leading the field, and others lagging behind in even basic infrastructure (Redecker & Punie, 2017).

Despite China's prominent position in AI worldwide and its ambitious policies in educational technology, there is a paucity of research on the factors of teacher AI literacy in schools, especially at the K-12 level (Hu et al., 2021). We hypothesize that there are substantial gaps between technological availability and classroom use (Li, Chen & Liu, 2023). By utilizing Bronfenbrenner's ecological systems theory as a theoretical lens, this study

Cite this article: Wu, X. and Annamalai, N. 2026. "Ecological Predictors of AI Literacy in Chinese K-12 Teachers: A Structural Equation Modeling Study", *Electronic Journal of e-Learning*, 24(2), pp 47-60, https://doi.org/10.34190/ejel.24.2.4155

explores how social environment and school factors interact with teacher self-efficacy to influence the development of AI literacy among Chinese teachers.

Research Objectives:

- To test the mediating effect of teacher self-efficacy in the relationship between school/social environment and AI literacy through SEM..
- To quantitatively describe current levels of AI literacy among Chinese K-12 teachers through different demographic characteristics.
- To identify and measure the relative impact of ecological factors (social support, school climate, self-efficacy) on AI literacy through validated instruments.

2. Literature Review

2.1 Definition of Key Concepts

2.1.1 AI literacy: A multidimensional construct

Artificial intelligence literacy is a new requirement for teachers with a background in intelligent education and critical literacy for teachers' competent education and teaching practice in the context of intelligent education (Sperling, et al., 2024). Compared with the traditional requirements of teachers' educational technology ability, the connotation of artificial intelligence literacy is more prosperous, and its requirements and challenges for teachers are also higher. According to related researches, AI literacy usually encompasses four dimensions:

- Awareness: Understanding AI concepts and societal implications (Long & Magerko, 2020).
- Application: Using AI tools for lesson design and student engagement (Guo & Wang, 2025).
- Evaluation: Critically assessing AI outputs for accuracy and bias (Hanna et al., 2024).
- Ethics: Navigating privacy, equity, and transparency challenges (Touretzky et al., 2019). For teachers, AI literacy transcends technical proficiency; it requires integrating tools like adaptive learning platforms while fostering ethical reasoning among students (Velander et al., 2023).

2.1.2 Ecological systems theory

Research on the influencing factors of AI literacy among K-12 teachers is relatively scarce. Among the existing research on the influencing factors of teachers' data and information literacy, mostly employing qualitative researches focus on the single prospective including the importance of information, the importance of learning digital skills and the importance of technology (Audrin, C., & Audrin, B, 2022). Bronfenbrenner's framework posits that teacher growth is shaped by nested systems:

- Macro: National policies, cultural attitudes, and technological infrastructure.
- Meso: School resources, leadership, and peer collaboration.
- Micro: Individual traits like self-efficacy and openness to innovation.
- Prior studies highlight school environments as catalysts for digital literacy (Powell & Bodur, 2019), while societal support (e.g., funding, public trust) sets the stage for scalable change (Yeager et al., 2022).

2.1.3 Teacher self-efficacy

Self-efficacy is a person's belief in their ability to perform a task, developed under social cognitive theory. Bandura's self-efficacy theory posits that teachers' confidence in using AI predicts adoption behaviors (An et al., 2023). High self-efficacy correlates with persistence in overcoming technical barriers and experimenting with AI tools. Despite having a low degree of self-efficacy, a teacher may do well on a given assignment (Compeau & Higgins, 1995). The same is true of teacher self-efficacy beliefs. A teacher may perform well on a specific task but may need a higher level of self-efficacy.

2.2 Relationships among the Key Concepts

2.2.1 Relationship between the competencies of teaching environments

Ecological teaching theory emphasizes that the learning environment profoundly shapes cognitive, emotional, and behavioral development by mediating the interactions between learners and their surrounding systems (Leijen, Pedaste & Lepp, 2020). This environment encompasses both the broader social context—such as societal norms, family dynamics, and cultural values—and the immediate school setting, including classrooms, pedagogical resources, and institutional structures. Within schools, elements like collaborative learning spaces and technology integration cultivate learners' social adaptability, ethical values, and interpersonal skills, thereby extending ripple effects beyond the classroom (Leijen, Pedaste & Lepp, 2020). In the era of artificial intelligence (AI), these dynamics evolve further: AI-enriched school environments—characterized by adaptive tools, ethical AI curricula, and interdisciplinary projects—can proactively reshape the social environment by equipping students with AI literacy that informs societal discourses on technology, equity, and innovation (Hartinah et al., 2020). Empirical evidence underscores this causal pathway, as AI-enhanced assessments and tutoring systems in schools promote social competencies such as teamwork, communication, and help-seeking behaviors, which in turn reduce inequities and normalize inclusive norms across communities (Bigman, 2025). By integrating AI to support culturally responsive pedagogies and human-centered designs, school environments thus serve as catalysts for broader social transformation. Based on this, the following hypothesis is proposed:

H1. School Teaching Environment has a positive effect on Social Environment.

2.2.2 Relationship between the competencies of teaching environment and perceived self-environment

In its realization, man's essence is the sum of all social relations. Marxist philosophy holds that sociality is the essential attribute of human beings. Human growth and development are bound to be restricted by the social environment. The formation of various qualities of K-12 teachers cannot exist independently from society, and the social environment significantly impacts teachers' cognition and behavior. As the cradle of young people's growth, primary and secondary schools are the main places to train successors for society. As one of the subjects of primary and secondary education activities, teachers' cognition and behavior are closely related to regional economic development and educational policies and regulations (Hanushek, Piopiunik & Wiederhold, 2019). Based on the above analysis, the following hypotheses are proposed in this study:

H2. School Teaching Environment has a positive effect on Teacher self-efficacy.

H3. Social Teaching Environment has a positive effect on Teacher self-efficacy.

2.2.3 Relationship between the competencies of teaching environment and AI literacy

According to ecosystem theory, the teaching environment can be divided into a macroscopic social teaching environment and a middle school teaching environment. In the existing studies on the impact of teaching on teacher literacy, social and cultural environment has a potential impact on teachers' academic literacy and information literacy, etc (Bury, 2016). The school environment positively impacts teachers' individual and professional development, among which campus cultural atmosphere, teacher relationship, teacher-student relationship, and training opportunities are essential factors affecting teachers' professional literacy (Powell & Bodur, 2019). Social environment (modified nature, interpersonal relationship, and social consciousness form), including the exceptional environment of the school, is an objective factor affecting the growth of teachers and plays a vital role in the cultivation and development of teachers' self-quality (Yeager, et al., 2022). Based on the above analysis, the following hypotheses are proposed in this study:

H4. School Teaching Environment has a positive effect on Teacher AI Literacy.

H5. Social Teaching Environment has a positive effect on Teacher AI Literacy.

2.2.4 Relationship between the competencies of perceived self-environment and AI literacy

K-12 teachers are the main objects of AI literacy development, and their self-efficacy in AI literacy awareness, technology application knowledge, intelligent evaluation ability, ethical cognition, and other aspects have an essential impact on the development of AI literacy (Lim, 2023). Teacher self-efficacy, as a subjective factor of teachers, determines the different levels and directions of teachers' individual growth in the same environment (Saglam, et al., 2023). Yang et al. pointed out that teachers' practical knowledge and perception of the ease of use of artificial intelligence affect teachers' judgment on the application value of artificial intelligence in education (Yang, Luo & Dong, 2020). In summary, teacher self-efficacy is believed to have a specific impact on teachers' artificial intelligence literacy. Based on this, the following hypothesis is proposed:

H6. Teacher self-efficacy has a positive effect on Teacher AI Literacy.

Based on the theoretical and conceptual review, a conceptual framework has been developed showing the relationship between the variables of the study as well as the interconnectedness of the hypotheses guiding the study (See Fig. 1).

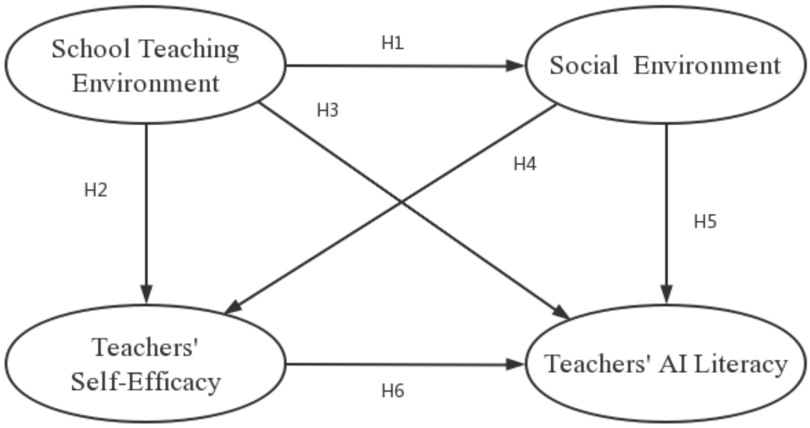


Figure 1: The hypothesized research model

3. Method

3.1 Sample

A four-month survey conducted from October 2022 to February 2023 is used for the study. A total of 1,970 questioners were collected randomly from formal K-12 schools in central China. The data was carefully screened through two quality control criteria: (1) Time validation of completion. Completion time should not be less than 120 seconds for participants to give adequate responses; (2) Completeness of responses. Questioners with over 50% of the values missing in the key sections of educational background and AI literacy items were eliminated as invalid. After careful screening, 1,680 valid samples were retained for structural equation modeling analysis (representing an 85.3% response rate).

Sampling was conducted in central China, which includes a large proportion of rural and suburban schools representative of Chinese educational settings. It is well known that central China has 60% of China's basic education schools. Sampling in urban China would result in a more biased sample of Chinese teachers and would inflate technological capabilities of teachers when compared to those from central China. Geographic distribution (47.3% rural, 19.6% suburban, 33.1% urban) reflects the geographic distribution of educational institutions in central China and is more generalizable to the broader Chinese population.

3.2 Participants

This study used a survey to explore the current state of AI literacy among Chinese K-12 teachers and identify related influencing factors. Teachers with at least one year of working experience in formal Chinese schools (grades 1-12) were eligible to participate. To reduce possible bias in sampling from different subject areas, we used systematic random sampling and proportionally allocated the sample according to the actual proportion of teachers from different subjects in the target areas.

All participants who agreed to participate in the survey provided informed consent, and the study was conducted in an ethical research according to the institutional review board. Demographic information listed in Table 1 indicates that although the proportion of females (52.8%) is slightly different from the reported proportion of Chinese K-12 teachers who are approximately 53% women (Hu, 2024), it is reasonable to assume that females are overrepresented in this sample. The elementary and secondary education level distribution is expected because China as an educational system has the majority of teachers at these levels. The hierarchical distribution of titles (34.6% beginner, 44.5% intermediate, 20.1% deputy senior, 0.8% senior) is reasonable because most schools in China have a pyramid structure in terms of professional title levels.

Table 1: Demographic information of the selected participants

Demographic	Level	Sample Size	Percentage	Total
Gender	Male	793	47.2%	1680
	Female	887	52.8%	
Location	Rural	794	47.3%	1680
	Suburban	330	19.6%	
	Urban	556	33.1%	
Educational Background	Below Bachelor	37	2.2%	1680
	Bachelor	1500	89.3%	
	Postgraduate	143	8.5%	
Age	≤31	302	33.1%	1680
	31-40	513	30.5%	
	41-50	591	35.2%	
	≥51	274	16.3%	
Title	Beginner	581	34.6%	1680
	Intermediate	747	44.5%	
	Deputy Senior	338	20.1%	
	Senior	14	0.8%	
Teaching Subject	Liberal Art	876	52.1%	1680
	Science	526	31.4%	
	Others	278	16.5%	

3.3 Instrument

This study adopted a culturally adapted questionnaire based on two published frameworks of AI literacy (Ng et al., 2021; Zhao, Wu & Luo, 2022). The adaptation was conducted in three systematic steps (Liu et al., 2022): (1) Translation and back-translation by bilingual education experts; (2) Content validation by expert panel (n=5 specialists in AI education); (3) Pilot testing with 150 teachers to validate construct validity in the Chinese cultural context. The final version of the questionnaire contained three parts: “Basic Information of K-12 teachers”, “Current Status of AIL among K-12 Teachers”, and “Influencing Factors of AIL among K-12 Teachers”. The AIL scale consisted of 20 items measuring four constructs: AI awareness literacy (KUAI), AI application literacy (AAI), AI evaluation literacy (EAIA), and AI ethics literacy (AIE). The influencing factors part was based on UNESCO’s Framework for Teachers’ ICT Competence (2018) and relevant literature and contained three parts: school teaching environment (STE, 4 items), social environment (SE, 5 items), and teachers’ self-efficacy (TSE, 5 items) (Hu, Zhang & Wang, 2021; Sanchez, 2013).

In addition to the original adaptation, CFs analysis were also conducted to verify the validity of construct characteristics in the Chinese context. The four-factor model of AI literacy presented a satisfactory fit (χ^2/df = 2.84, CFI = 0.934, TLI = 0.921, RMSEA = 0.052, SRMR = 0.048), indicating that the adapted instrument was suitable for use on AI literacy of Chinese K-12 teachers. All items used a 5-point Likert scale (1 = “strongly disagree” to 5 = “strongly agree”), and a higher score indicated stronger agreement or competency.

Internal consistency was assessed using Cronbach’s α coefficients, and excellent reliability were obtained for AI literacy scale (α = 0.937) and good reliability were obtained for influencing factors scale (α = 0.875). Composite reliability for each constructed was calculated and ranged from 0.841 to 0.923, all larger than the recommended level of 0.7. Construct validity of the measurement was assessed through exploratory (EFA) and confirmatory factor analysis (CFA). EFA results showed that KMO was 0.937 (>0.7) and the Bartlett’s sphericity test was highly significant ($p < 0.001$), indicating good sampling adequacy. CFA results confirmed the expected factor structure. All factor loadings were larger than 0.6 and average variance extracted (AVE) values ranged from 0.578 to 0.694, which all met the convergent validity criteria.

3.4 Data Analysis

Two analytical strategies were employed. Descriptive and correlation analyses were used to describe the levels of AI literacy and explore associations between measurement variables and demographic variables (gender, age, educational background, teaching level, location). Structural equation modeling (SEM) is particularly good at maximizing predictive power (Hair et al., 2019), thus it could help this research optimal for exploring factors of teachers' AIL. SPSS 22.0 was used for descriptive statistics and AMOS 24.0 for SEM analysis.

4. Results

4.1 Descriptive and Correlational Statistics

Table 2 presents the means and standard deviations for teachers' AI literacy (AIL), school teaching environment (STE), social teaching environment (SE), and teachers' self-efficacy (TSE). Following recommendations for structural equation modeling (Hu & Bentler, 1999), we analyze the basic statistical assumptions of structures for such data sets. The mean scores for AIL (M = 4.029), STE (M = 4.281), SE (M = 4.744), and TSE (M = 4.245) all exceeded the midpoint of 3 on a 5-point Likert scale, indicating generally positive perceptions among teachers. These above-average scores suggest teachers possess relatively high views of their AI literacy, perceive supportive teaching environments, and exhibit strong self-efficacy, which may facilitate the adoption of AI in educational settings.

Table 2: Mean and Standard deviation of the key constructs

	Items	Mean(SD)
AIL	20	4.029 (0.862)
STE	5	4.281(0.828)
SE	5	4.744 (0.763)
TSE	5	4.245(0.798)

Table 3 displays correlations among key constructs and demographics. Notably, AIL showed strong positive correlations with STE (r = 0.609, p < 0.01, 95% CI [0.57, 0.65]), SE (r = 0.355, p < 0.01, 95% CI [0.31, 0.41]), and TSE (r = 0.637, p < 0.01, 95% CI [0.60, 0.68]), suggesting that teachers' AI literacy is closely tied to their teaching environments and self-efficacy (Cohen's d ≈ 0.5–0.8 for practical significance). These correlations between AIL, STE, SE, and TSE highlight their interconnectedness, where gains in one (e.g., school environment) may enhance others, though effect sizes imply context-specific influences rather than universal drivers.

Table 3: Correlations of the key constructs

	D1	D2	D3	D4	D5	D6	AIL	STE	SE	TSE
D1	1									
D2	-0.311**	1								
D3	-0.198**	0.053*	1							
D4	0.002	0.000	-0.005	1						
D5	-0.135**	0.290**	0.052*	0.064**	1					
D6	-0.297**	0.711**	-0.017	-0.061*	0.193**	1				
AIL	0.036	-0.055*	0.007	-0.010	-0.030	-0.020	1			
STE	0.022	0.015	0.003	0.549**	-0.004	0.006	0.609**	1		
SE	0.016	-0.008	-0.002	0.019	0.007	-0.032	0.355**	0.497**	1	
TSE	0.035	0.446*	0.012*	0.528**	0.134**	-0.012	0.637**	0.557**	0.552**	1

Note: ** p < 0.01; * p < 0.05. D1, Gender; D2, Age; D3, Teaching Subject; D4, Location; D5, Educational Background; D6, Ttitle; SD, standard deviation; AIL, AI Literacy; STE, School Teaching Environment; SE, Social Environment; TSE, Teachers' Self-Efficacy.

4.2 Measurement Model

Reliability and validity of the questionnaire were assessed to ensure the consistency and accuracy of the measurement tools. Reliability, reflecting the stability of test results, was evaluated using Cronbach's α via SPSS.

According to Nunnally (1978), α values above 0.7 indicate high reliability. Validity, which measures the accuracy of the instrument, was assessed using factor loadings and the Kaiser-Meyer-Olkin (KMO) test, with KMO values above 0.7 considered acceptable (Sanchez, 2013). Table 4 summarizes metrics: Cronbach's α (0.84–0.87), composite reliability (CR; 0.93–0.96), and average variance extracted (AVE; 0.64–0.76), all meeting thresholds. Fornell-Larcker criteria confirmed discriminant validity (square root of AVE > inter-construct correlations). These results indicate that the questionnaire reliably and accurately measures AIL, STE, SE, and TSE, providing a solid foundation for the structural model analysis.

Table 4: Statistical summary of reliability and validity of the questionnaire

	Items	Cronbach's α	Factor Loading	CR	AVE
AIL	20	0.843	0.807–0.867	0.927	0.724
STE	5	0.871	0.737-0.851	0.938	0.643
SE	5	0.836	0.835–0.924	0.942	0.756
TSE	5	0.859	0.825–0.905	0.961	0.693

4.3 Structural Model

Figure 2 shows the interaction between demographic variables, teacher AI literacy (AIL), school teaching environment (STE), social teaching environment (SE), and teacher self-efficacy (TSE) and their mutual influence, with significant SEM results inserted. Model fit was assessed using multiple indices, all indicating a good fit: CFI = 0.986 (>0.900), RMSEA = 0.030 (<0.080), SRMR = 0.062 (<0.080), and Chi-square test ($p < 0.001$), as shown in Table 5, aligning with Hair et al. (2017).

Table 5: Results of the fitness in the hypothesis model

	P	CFI	RMSEA	SRMR
Structural Model	0.000	0.986	0.030	0.062
Fit Criteria	＜0.001	＞0.900	＜0.080	＜0.080

Non-significant pathways were removed, and significant path coefficients were retained in Table 6. Significant paths included STE to SE (β = 0.245, $p < 0.001$), STE to TSE (β = 0.295, $p < 0.001$), SE to TSE (β = 0.221, $p < 0.001$), STE to AIL (β = 0.270, $p < 0.001$), and TSE to AIL (β = 0.259, $p < 0.001$). The path from SE to AIL (β = 0.060, p = 0.362) was non-significant, rejecting hypothesis H5. Paths suggest STE and TSE as primary AIL drivers (total R^2 = 0.45, moderate effect), with SE indirectly influencing via TSE (indirect β = 0.05, 95% CI [0.01, 0.10]). This aligns with reciprocal ecological influences, where school-level factors mutually shape social contexts without strict causality..

Table 6: Path coefficient estimates in the hypothesized model

	Path Coefficient (β)	Direct Effects	P
STE--->SE	0.245	0.065	0.000
STE--->TSE	0.295	0.061	0.000
SE--->TSE	0.221	0.063	0.000
STE--->AIL	0.270	0.081	0.000
SE--->AIL	0.060	0.066	0.362
TSE--->AIL	0.259	0.079	0.000

Note: *** $p < 0.001$. AIL, AI Literacy; STE, School Teaching Environment; SE, Social Environment; TSE, Teachers' Self-Efficacy.

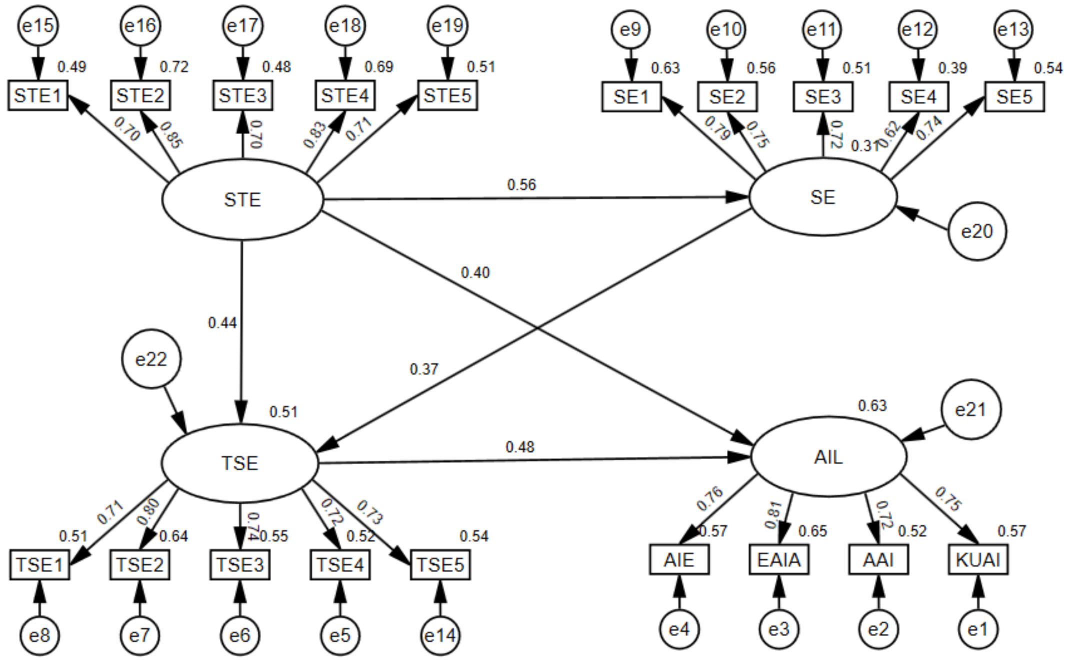


Figure 2: Structural model of the relationships among School Teaching Environment, Social Environment, Teachers' Self-Efficacy and AI Literacy

5. Discussion

5.1 Bridging the AI Application-Evaluation Gap

The results indicate that Chinese K-12 teachers have moderate levels of AI literacy (M = 3.89 out of 5), and almost half of the teachers scored below the mean. More worryingly, the theory-practice gap was significant: stronger performance in the theoretical dimension—awareness (KUAI, $\beta = 0.75$) and ethics (AIE, $\beta = 0.76$), but weaker performance in practical competencies—application literacy (AAI, $\beta = 0.72$) and evaluation literacy (EAIA, $\beta = 0.81$). Teachers grasp AI concepts but show modest operational skills (effect size $d \approx 0.3$–0.5), limiting classroom integration.

This finding contrasts markedly with international research. Systematic government-led AI integration initiatives, along with strong technical infrastructure, result in higher application competencies among teachers in Nordic countries (Saheb & Saheb, 2022). Likewise, comprehensive professional development programs focusing on hands-on practice at institutional and systemic levels result in higher application competencies among Singaporean educators (Chai et al., 2023). The substantial gap between low AI literacy and fast technological development offers increasingly great challenges for educational digitalization in China and might widen existing digital divides in Chinese education.

5.2 School Environments as Catalysts

School teaching environment (STE) was found to be the strongest predictor of AI literacy ($\beta = 0.270$, $p < 0.001$) and Hypothesis 1 was supported. In agreement with Bronfenbrenner's ecological systems theory, microsystem environments, which are the immediate settings where individuals are embedded and involved in direct interactions, have the primary influence on developmental processes (Krebs, 2009). As a proximal environment, schools have the most immediate and longest influence on teachers' technological competency development.

Demographic analyses provide insights into the ecological nature of teachers' AI literacy development. Gender differences show female teachers exhibiting higher AI literacy than their male counterparts, contrasting with Western research indicating male advantage in technology use (Steinberg & Hohenberger, 2023). This may reflect the feminized nature of China's teaching profession and women's enhanced collaborative learning orientations. Age-related patterns demonstrate that younger teachers (under 30) exhibit the highest AI literacy levels, with sharp declines observed after age 40. This supports ecological systems theory's emphasis on environmental resource availability and suggests generational digital divide effects (Lim, 2023).

Geographic disparities reveal significant urban-rural differences in AI literacy, reflecting what we term "digital ecosystem disparities." Urban schools typically possess richer technological microsystems and superior access to macrosystem policies. Subject-based variations (science teachers outperforming liberal arts teachers) indicate that disciplinary cultures influence technology adoption, with STEM fields potentially establishing more favorable mesosystem connections to AI technologies.

5.3 Self-efficacy: The Linchpin

Teacher self-efficacy (TSE) shows partial mediating effects ($\beta = 0.259$, $p < 0.001$), which confirm Hypothesis 3. Social cognitive theory holds self-efficacy is an important factor in behavioral change (Bandura, 1986). Teachers with high self-efficacy tend to see AI as an opportunity rather than a threat and display exploratory behaviors to improve their competencies (Hartinah et al., 2020). School exerts influence on AI literacy directly ($\beta = 0.270$) and indirectly through self-efficacy ($\beta = 0.295 \times 0.259 = 0.076$), indicating partial mediation in which schools affect teachers' AI literacy through two pathways: resource provision and confidence building.

Schools can improve teacher self-efficacy by implementing interventions reflecting Bandura's four sources of self-efficacy: (1) hierarchical teacher training programs providing mastery experiences (Martins & Von Wangenheim, 2022); (2) peer mentoring systems offering teachers modeling and vicarious experiences along with emotional support (Alam, 2021); (3) administrative recognition and social persuasion towards teachers' AI integration practices; and (4) comprehensive technical support systems reducing implementation and facilitate emotional regulation (Street, 2013).

5.4 The Non-significant Social Pathway

The non-significant direct effect of social environment on AI literacy ($\beta = 0.066$, $p = 0.362$) carries theoretical significance as it challenges assumptions regarding policy effectiveness, failing to support Hypothesis 5. This finding suggests that distal policy interventions may prove ineffective without translation through school-level mediations, aligning with Bronfenbrenner's mesosystem emphasis. It implies that macrosystem variables—including national AI policies, media discourse, and public attitudes—may insufficient to influence teacher competencies without proximal mediation mechanisms.

This pattern may reflect implementation gaps in China's AI education policies. While the State Council's "AI Development Plan" (2017) established clear objectives for teacher training in AI education, resource allocation and support mechanisms vary considerably across regions. Additionally, social environment effects may operate on longer time scales than captured in this cross-sectional study, as policies possess inherent lifecycles requiring years to reach educational practice. Within Chinese educational contexts, social influences may traverse multiple administrative hierarchies, diminishing direct effects while strengthening indirect pathways through school leadership (Sui et al., 2020; Vestrucci, Lumbreras & Oviedo, 2021).

The findings strongly support Bronfenbrenner's (2000) ecological systems theory, suggesting that connections between microsystems (school-community linkages) may be more meaningful than isolated macrosystem effects. The indirect effect of social environment through school environment ($\beta = 0.245 \times 0.270 = 0.066$) indicates that when social policies fail to reach schools effectively, their intended impacts may never materialize. This contrasts with studies in federalized systems such as Germany or Australia, where direct effects were observed (Niemann, Eickelmann & Drossel, 2025), or Finland, where indirect effects successfully bridged macro and micro levels through robust school-community partnerships (Malone, 2020).

5.5 Theoretical Integration Within Ecological Systems Framework

Ecological systems theory posits that microsystem environments are primary environments for development. Schools are immediate and prolonged interaction environments where teachers obtain resources, training, and peer support for development of AI literacy. Significant positive correlations between school and social environments ($\beta = 0.245$) show that effective AI literacy development requires connection at the meso level of ecological systems, including school-based initiatives that connect with broader social support systems such as policy contexts, community resources, and peer networks.

Schools have the indirect effect of social environment on teacher development through school environment ($\beta = 0.066$) shows that macrosystem variables such as culture and values, policy, and technology do influence conditions at the school level where teachers are developed indirectly. Findings on demographic patterns show that historical timing of technology use leads to current competencies. Cohort effects (age $r = -0.06$) urge stage-tailored support, integrating ecological layers for holistic AI literacy growth.

6. Conclusion and Implications

In this study, structural equation modeling was used to explore ecological predictors of AI literacy among Chinese K-12 teachers. Findings show moderate levels of AI literacy (M = 3.89 out of 5), with a theory-practice gap: teachers display stronger performance in AI awareness ($\beta = 0.75$) and ethics ($\beta = 0.76$) but significant weaknesses in practical application ($\beta = 0.72$) and evaluation skills ($\beta = 0.81$). The ecological analysis shows that school teaching environment exerts the strongest direct effects on AI literacy ($\beta = 0.270$, $p < 0.01$), while teacher self-efficacy is a significant mediator ($\beta = 0.259$, $p < 0.01$). Notably, social environment shows no significant direct effects ($\beta = 0.066$, $p = 0.362$), suggesting that distal policy influences require proximal school-level mediation to impact teacher competencies.

6.1 Implications for Government and Policy Makers

Given the non-significant direct effect of social environment, policy makers should focus less on developing macro-level policies and more on providing concrete school-level support by providing funding for school infrastructure and training programmes. Considering the steep decline in AI literacy after the age of 40, cohort-specific interventions should be provided for different age groups, including mentorship systems and age-appropriate professional development programmes.

Policy recommendations are as follows: (1) build regional AI education resource centres to assist schools in implementing programmes; (2) design differentiated funding policies to address urban-rural divides; (3) construct cohort and career-stage-specific professional development frameworks; and (4) build policy monitoring systems that focus on school-level outcomes instead of system-level inputs.

6.2 Implications for School Leaders

Given that school is one of the significant predictors of AI literacy ($\beta = 0.270$), school leaders play a critical role in shaping enabling environments for AI integration. However, considering that not all schools have the infrastructure, funding, or policy support to implement AI technologies immediately. Therefore, school leaders should adopt a phased and context-sensitive approach when promoting AI literacy. Schools with greater resources may gradually introduce designated AI learning spaces, structured collaboration opportunities, and technical support systems. School leaders should implement programmes that address the four sources of Bandura's model of self-efficacy, including scaffolded mastery experiences, peer modelling opportunities, administrative encouragement and stress-reduction activities: (1) developing school-based AI integration strategic plans; (2) peer modelling through classroom observations and lesson planning; (3) establishing technical support protocols for AI tool implementation; (4) developing stress-reduction activities to mitigate technology-related anxiety.

6.3 Implications for Teachers

As teachers are the strongest predictors of AI literacy ($\beta = 0.72$ and $\beta = 0.81$), they should actively engage in AI literacy enhancement and maximise self-efficacy development. Recognizing barriers to advanced AI tool access, emphasis should be placed on foundational AI knowledge, pedagogical adaptations using accessible open-source resources, and self-directed growth. To bridge competency gaps, teachers can prioritize: (1) engaging in experiential AI training via open-access tools or platforms; (2) constructing practical and context-appropriate criteria to evaluate AI-generated content; (3) joining peer collaboration networks to share insights and teaching practices; (4) exploring subject-specific integration strategies that align with available resources; (5) cultivating a growth mindset to stay informed about evolving AI developments.

7. Limitations and Future Research

This study is subject to two principal limitations. First, the sample was predominantly drawn from a central Chinese province—strategically selected as a "middle" region to approximate national averages in socioeconomic development and educational infrastructure, thereby capturing a representative midpoint in AI literacy distribution. While this purposive choice mitigated extremes associated with coastal metropolises or remote western areas, the geographic concentration nonetheless constrains generalizability to China's heterogeneous regional contexts, including more affluent eastern provinces (e.g., Guangdong) or agrarian western ones (e.g., Gansu). Compounding this, the non-random sampling further tempers extrapolative claims. Future research should prioritize probability-based sampling across multifaceted strata, encompassing diverse provinces, urban-rural locales, and institutional archetypes, to bolster external validity. Second, the reliance on cross-sectional quantitative data precludes causal inferences or temporal dynamics. Subsequent studies should integrate mixed-methods approaches, including qualitative explorations of lived experiences and longitudinal tracking of AI literacy evolution, to illuminate developmental pathways. Notwithstanding these constraints, the substantive findings retain robustness within the sampled context.

AI Statement: No AI tools were used in the preparation of this manuscript. All aspects of the work were completed solely by the authors, who take full responsibility for the originality, accuracy, and integrity of the content.

Ethics Statement: All subjects gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of Ajman University (Reference number: H-F-H-3—A) on 31 August 2023.

References

Alam, A. (2021). Possibilities and apprehensions in the landscape of artificial intelligence in education. In Proceedings of the 2021 International Conference on Computational Intelligence and Computing Applications (ICCICA) (pp. 1-8). IEEE.

An, X., Chai, C. S., Li, Y., Zhou, Y., Shen, X., Zheng, C., & Chen, M. (2023). Modeling English teachers' behavioral intention to use artificial intelligence in middle schools. Education and Information Technologies, 28(3), 5187-5208.

Audrin, C., & Audrin, B. (2022). Key factors in digital literacy in learning and education: A systematic literature review using text mining. Education and Information Technologies, 27(6), 7395-7419.

Bigman, M. (2025). Pear Program: Leveraging AI and the Learning Sciences to Create More Fruitful Collaboration (Doctoral dissertation, Stanford University).

Bronfenbrenner, U. (2000). Ecological systems theory. American Psychological Association.

Bury, S. (2016). Learning from faculty voices on information literacy: Opportunities and challenges for undergraduate information literacy education. Reference Services Review, 44(3), 237-252.

Chai, C. S., Lin, P. Y., Jong, M. S. Y., Dai, Y., Chiu, T. K., & Qin, J. (2021). Perceptions of and behavioral intentions towards learning artificial intelligence in primary school students. Educational Technology & Society, 24(3), 89-101.

Compeau, D. R., & Higgins, C. A. (1995). Computer self-efficacy: Development of a measure and initial test. MIS Quarterly, 19(2), 189-211.

Education Week Research Center. (2023). Teachers and artificial intelligence: National survey results. Education Week.

Guo, Y., & Wang, Y. (2025). Exploring the effects of artificial intelligence application on EFL students' academic engagement and emotional experiences: A mixed-methods study. European Journal of Education, 60(1), e12812.

Hair, J. F., Jr., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2019). A primer on partial least squares structural equation modeling (PLS-SEM) (2nd ed.). Sage Publications.

Hair, J. F., Jr., Matthews, L. M., Matthews, R. L., & Sarstedt, M. (2017). PLS-SEM or CB-SEM: Updated guidelines on which method to use. International Journal of Multivariate Data Analysis, 1(2), 107-123.

Hanna, M., Pantanowitz, L., Jackson, B., Palmer, O., Visweswaran, S., Pantanowitz, J., & Rashidi, H. (2024). Ethical and bias considerations in artificial intelligence/machine learning. Modern Pathology, 37(1), 100686.

Hanushek, E. A., Piopiunik, M., & Wiederhold, S. (2019). The value of smarter teachers: International evidence on teacher cognitive skills and student performance. Journal of Human Resources, 54(4), 857-899.

Hartinah, S., Suharso, P., Umam, R., Syazali, M., Lestari, B. D., Roslina, R., & Jermsittiparsert, K. (2020). Retracted: Teacher's performance management: The role of principal's leadership, work environment and motivation in Tegal City, Indonesia. Management Science Letters, 10(4), 883-892.

Hu, L. (2024). Exploring gender differences in computational thinking among K-12 students: A meta-analysis investigating influential factors. Journal of Educational Computing Research, 62(5), 1138-1164.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling, 6(1), 1-55.

Hu, X., Yuan, S., Xu, K., et al. (2021). Analysis of factors influencing the application literacy of intelligent technology for K-12 teachers. Education Information Technology, 36(5), 1-15.

Hu, X., Zhang, J., & Wang, L. (2021). Artificial intelligence in Chinese education: Current status and future prospects. Educational Technology Research and Development, 69(4), 1951-1970.

Krebs, R. J. (2009). Proximal processes as the primary engines of development. International Journal of Sport Psychology, 40(1), 219.

Leijen, Ä., Pedaste, M., & Lepp, L. (2020). Teacher agency following the ecological model: How it is achieved and how it could be strengthened by different types of reflection. British Journal of Educational Studies, 68(3), 295-310.

Li, S., Chen, Y., & Liu, H. (2023). The gap between policy and practice: AI integration in Chinese K-12 education. Computers & Education, 195, 104720.

Lim, E. M. (2023). The effects of pre-service early childhood teachers' digital literacy and self-efficacy on their perception of AI education for young children. Education and Information Technologies, 28(10), 12969-12995.

Liu, Z., Ren, Y., Kong, X., & Liu, S. (2022). Learning analytics based on wearable devices: A systematic literature review from 2011 to 2021. Journal of Educational Computing Research, 60(6), 1514-1557.

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 (pp. 1-16).

Malone, H. J. (2020). Community schools: Bridging educational change through partnerships. Journal of Educational Change, 21(3), 487-497.
Martins, R. M., & Von Wangenheim, C. G. (2022). Findings on teaching machine learning in high school: A ten-year systematic literature review. Informatics in Education, 21(1), 1-21.
Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. Computers and Education: Artificial Intelligence, 2, 100041.
Niemann, J., Eickelmann, B., & Drossel, K. (2025). Overcoming Digital Inequalities—Identification and Characterisation of Digitally Resilient Schools in Different Countries Using ICILS 2023 Data. Education Sciences, 15(7), 898.
Nunnally, J. C. (1978). Psychometric theory. American Journal of Educational Research, 15(3), 245-278.
Pedro, F., Subosa, M., Rivas, A., & Valverde, P. (2019). Artificial intelligence in education: Challenges and opportunities for sustainable development. UNESCO.
Powell, C. G., & Bodur, Y. (2019). Teachers' perceptions of an online professional development experience: Implications for a design and implementation framework. Teaching and Teacher Education, 86, 102929.
Redecker, C., & Punie, Y. (2017). European framework for the digital competence of educators: DigCompEdu. Publications Office of the European Union.
Saglam, M. H., Goktenturk, T., Demir, I., & Yazıcı, E. (2023). Environmental factors for the advancement of teachers' self-efficacy in professional development. Journal of Intelligence, 11(2), 28.
Saheb, T., & Saheb, T. (2023). Topical review of artificial intelligence national policies: A mixed method analysis. Technology in Society, 74, 102316.
Sanchez, G. (2013). PLS path modelling with R. Trowchez Editions.
Steinberg, A. L. S., & Hohenberger, C. (2023). Can AI close the gender gap in the job market? Individuals' preferences for AI evaluations. Computers in Human Behavior Reports, 10, 100287.
Sperling, K., Stenberg, C. J., McGrath, C., Åkerfeldt, A., Heintz, F., & Stenliden, L. (2024). In search of artificial intelligence (AI) literacy in teacher education: A scoping review. Computers and Education Open, 5, 100169.
Street, B. (2013). Literacy in theory and practice: Challenges and debates over 50 years. Theory into Practice, 52(1), 4-9.
Su, J., Zhong, Y., & Ng, D. T. K. (2022). A meta-review of literature on educational approaches for teaching AI at the K-12 levels in the Asia-Pacific region. Computers and Education: Artificial Intelligence, 3, 100065.
Sui, X., Zhao, G., Wang, J., et al. (2020). An empirical study on the factors influencing the informationization teaching ability of university teachers: Taking some universities in Hunan Province as an example. China's Audiovisual Education, 5, 18-25.
Touretzky, D., Gardner-McCune, C., Breazeal, C., et al. (2019). A year in K-12 AI education. AI Magazine, 40(4), 88-90.
Velander, J., Taiye, M. A., Otero, N., et al. (2023). Artificial intelligence in K-12 education: Eliciting and reflecting on Swedish teachers' understanding of AI and its implications for teaching & learning. Education and Information Technologies, 28(7), 7671-7701.
Vestrucci, A., Lumbreras, S., & Oviedo, L. (2021). Can AI help us to understand belief? Sources, advances, limits, and future directions. International Journal of Interactive Multimedia and Artificial Intelligence, 6(7), 7-14.
Yang, Y., Luo, Z., & Dong, Y. (2020). Future teacher's work innovation, cross-border and reshaping. Open Education Research, 26(1), 27-35.
Yeager, D. S., Hanselman, P., Walton, G. M., Murray, J. S., Crosnoe, R., Muller, C., Tipton, E., Schneider, B., Hulleman, C. S., Hinojosa, C. P., Paunesku, D., Romero, C., Flint, K., Roberts, A., Trott, J., Iachan, R., Buontempo, J., Yang, S. M., Carvalho, C. M., & Dweck, C. S. (2022). Teacher mindsets help explain where a growth-mindset intervention does and doesn't work. Psychological Science, 33(1), 18-32.
Zhao, L., Wu, X., & Luo, H. (2022). Developing AI literacy for primary and middle school teachers in China: Based on a structural equation modeling analysis. Sustainability, 14(21), 14549.

Appendix 1: Questionnaire

Basic information

1. Your gender is

□Male □Female

2. Your school location is

□rural

□suburban

□urban

3.Your educational background is

□Below Bachelor

□Bachelor

□Postgraduate

4. Your age is

□≤30 years old

□31–40 years old

□41–50 years old

□≥51 years old

5.Your title is

□Beginner

□Intermediate

□Deputy Senior

□Senior

6.Your teaching subject is

□Liberal Art

□Science

□Others

Knowing and Understanding AI

1. I am able to discern between smart and non-smart instructional equipment.

2. I am aware of how AI in education may benefit me.

3. I can recognise AI technology in goods or services that are intended for use in education.

4. When utilising educational AI goods, I feel at ease.

5. In Smart era, I believe that educators should actively learn how to employ sophisticated technology to support learning and teaching.

Applying AI

1. I am competent in using educational AI tools to assist me in my day-to-day teaching activities.

2. I can pick up new educational AI products rather quickly.

3. I can increase the effectiveness of my instruction by using educational AI solutions.

4. I may assist and direct pupils as they use instructional AI products.

5. I can use curriculum teaching and educational AI technologies.

Evaluating AI Application

1. I am proficient in utilising educational AI technologies to help me with my daily teaching tasks.

2. I am quite fast to take up new educational AI goods.

3. By using educational AI technologies, I can improve the efficacy of my lesson.

4. I may guide and help students while they utilise educational AI goods.

5. I am capable of using curriculum teaching and AI technology in education.

AI Ethics

1. When using educational AI products, I always abide by ethical standards.

2. I pay close attention to privacy, information security, and other problems while utilising educational AI products.

3. I'm concerned about the abuse of educational AI.

4. When implementing instructional AI technology, I constantly take safety and ethical concerns into account.

5. I have the ability to quickly spot ethical lapses in the use of artificial intelligence in the classroom.

School Teaching Environment

1. The school can meet the requirements of intelligent teaching hardware facilities (such as touch screen electronic whiteboard, interactive all-in-one computer, etc.).

2. The school can meet the requirements of intelligent teaching related software.

3. The school attaches great importance to the application of intelligent technology in teachers' teaching.

4. The school provides teachers with training opportunities to improve their intelligent educational literacy.

Social Environment

1. My province has clearly proposed policies or action plans related to teachers' intelligent educational literacy.

2. The district where I teach gives full financial support to the application of intelligent technology in teaching.

3. The region where I teach gives full publicity and support to the application of intelligent technology in teaching.

4. My district often holds events (such as smart technology teaching competitions, etc.) to encourage teachers to use smart technology in teaching.

5. My region often invites AI experts to conduct training activities.

Teacher Self-Efficacy

1. The application of intelligent technology in daily teaching will make the teaching effect double with half the effort.

2. Applying smart technology to everyday teaching will lead to better career development for me.

3. I am confident to choose the appropriate intelligent technology for course teaching according to the new curriculum standards.

4. By choosing the right smart technology, I can effectively promote student learning and improve the effectiveness of classroom teaching.

5. I think my own intelligent education quality can be effectively improved through teacher training and other ways.

Learning Design and Learning Analytics to Improve Higher Education: A Systematic Literature Review

Rangana Jayashanka, E. Hettiarachchi, Prasad Wilagama and K.P.Hewagamage
University of Colombo School of Computing, Sri Lanka

rja@ucsc.cmb.ac.lk
eno@ucsc.cmb.ac.lk
pdw@ucsc.cmb.ac.lk
kph@ucsc.cmb.ac.lk

https://doi.org/10.34190/ejel.24.2.3645

Abstract: In recent years, higher education has increasingly emphasized the integration of Learning Design and Learning Analytics to foster more engaging, personalized, and effective learning environments. This systematic literature review investigates how these two domains interact to enhance teaching learning processes and improve educational outcomes. The review identifies key benefits and opportunities associated with this integration across three stakeholder groups: students, lecturers, and educational institutions by analyzing 55 peer-reviewed publications. The results show that learning effectiveness can be significantly enhanced through the visualization of students' learning interactions using straightforward and user-friendly analytical approaches. Furthermore, successful implementation requires the development of lecturers' data literacy and programming competencies, as well as the incorporation of sociocultural, psychological, and physical data to achieve a more holistic understanding of learners. The review also identifies four major research directions to guide future efforts in bridging Learning Analytics and Learning Design. Finally, the paper underscores the importance of establishing clear ethical and privacy frameworks to ensure the responsible application of Learning Analytics in higher education.

Keywords: Learning analytics, Learning design, Technology-enhanced learning, Higher education, Learning analytics dashboards

1. Introduction

Learning is a process that changes how individuals understand themselves and the world around them (Looney and Siemens 2011). It allows students to gain new knowledge, improve their skills, and adapt their behaviors. Students participate in both formal and informal learning. Informal learning is often self-directed, peer-supported, and rooted in everyday practice, producing outcomes that can be flexible or unexpected (Nicolae, Mihai and Stefan 2019). In contrast, formal learning takes place in schools, universities, and other educational institutions, guided by structured curricula, defined goals, and formal assessments (Nicolae, Mihai and Stefan 2019).

In universities, tools like wikis, blogs, video conferencing systems, discussion forums, and MOOCs are commonly used to promote active learning. While these tools offer significant benefits, challenges such as low pass rates or unsatisfactory learning outcomes sometimes arise. One major reason is that the design of learning activities, commonly referred to as Learning Design (Craft and Mor 2012), may not align well with the intended learning outcomes. Lecturers often lack direct feedback on how students engage with these activities, making it difficult to refine and improve Learning Design in future iterations.

Learning Analytics provides a solution by analyzing student data to offer insights into engagement and performance. Learning Management Systems can collect detailed records of students' activities, which Learning Analytics can process to generate actionable insights in real time. By combining Learning Design and Learning Analytics, lecturers can better understand student behaviors, evaluate the effectiveness of their learning activities, and make evidence-based improvements.

This study focuses on identifying the benefits of combining Learning Design and Learning Analytics and highlighting opportunities to improve higher education through better assessment of learning activities and visualization of student's engagements.

This paper is structured as follows. Section 2 reviews the theoretical background of Learning Design and Learning Analytics, as well as their interrelationship. Section 3 outlines the research questions and describes the methodology employed in the study. Section 4 presents the results of the literature review, and Section 5 discusses the findings and offers concluding remarks.

Cite this article: Jayashanka, R. et al. 2026. "Learning Design and Learning Analytics to Improve Higher Education: A Systematic Literature Review", *Electronic Journal of e-Learning*, 24(2), pp 61-81, https://doi.org/10.34190/ejel.24.2.3645

2. Background

Both Learning Design and Learning Analytics aim to enhance the learning environments and improve the effectiveness of the learning process. Sharing a common aim increases the potential to have a synergy between these two domains (Mangaroska and Giannakos 2017). On the one hand, the outputs generated by well-formulated Learning Analytics tools can provide valuable information to lecturers and learning designers about the outcomes, success, and effectiveness of their Learning Designs (Alhadad and Thompson 2017; Wise et al. 2016). On the other hand, Learning Designs present domain vocabulary to represent the elements in an educational system where Learning Analytics can be applied. The natural and synergistic relationship that emerges between Learning Analytics and Learning Design has therefore led to the emergence of this fast-growing research area. Different definitions for Learning Design can be found in the existing literature. Ifenthaler and Gibson (2018) extended Dobozy's (2013) Learning Design definition list to further illustrate the roots of Learning Design by including the recent definitional constructs. The most comprehensive and most cited definitions are presented in **Table 1.**

Table 1: Learning Design Definitions

Author(s)	Definition
Agostinho (2006)	*"A learning design is a representative of teaching and learning practice documented in some notational form so that it can serve as a model or template adaptable by a teacher to suit his/her context."*
Conole (2013)	*"A methodology for enabling teachers/designers to make more informed decisions in how they go about designing learning activities and interventions, which is pedagogically informed and makes effective use of appropriate resources and technologies."*
Emin-Martinez et al. (2014)	*"The act of devising new practices, plans of activity, resources and tools aimed at achieving particular educational aims in a given situation."*
Koper (2006)	*"The description of the teaching-learning process that takes place in a unit of learning. The key principle in learning design is that it represents the learning activities and the support activities that are performed by different persons (learners, teachers) in the context of a unit of learning. These activities can refer to different learning objects that are used during the performance of the activities (e.g. books, articles, software programmes, pictures), and it can refer to services (e.g.: forums, chats, wiki's) that are used to collaborate and to communicate in the teaching-learning process."*
Hale (2016)	*"Learning Design is the process of designing learning experiences (planning, structuring, and sequencing) through facilitated activities that are pedagogically informed, explicit, and make better use of technologies in teaching."*

As a summary, Learning Design describes the learning activities, resources required, and the different ways a lecturer can use these to facilitate the learning. These activities are intended to increase student knowledge and skills through interaction with the lecturer, peers, or content items (Marcel et al. 2017). Furthermore, a Learning Design can be reused and customized by other lecturers since Learning Design activities are independent of its implementation context (Dalziel et al. 2016; Hernández-Leo et al. 2014). Most of the research work in the Learning Design field has focused on the creation of tools, practices, and processes, and on sharing their outputs with practitioners. Only a few studies have evaluated the effectiveness of Learning Designs. This highlights the need for methodological approaches capable of assessing learning effectiveness and identifying design elements that need redesign in subsequent iterations, a process in which Learning Analytics can play a critical supporting role.

This underscores the need for methodological approaches capable of assessing learning effectiveness and identifying design elements requiring refinement in subsequent iterations, a process in which Learning Analytics can play a critical supporting role.

Lecturers can use Learning Analytics to evaluate the success and outcomes of their Learning Designs (Lockyer, Heathcote, and Dawson 2013; Melero et al. 2015). Learning Design alone does not provide information on how students engage with the design during and after the course. Learning Analytics fills this gap by providing data during different stages of the course, enabling evaluation of the design and offering a holistic view of its impact (Lockyer, Heathcote, and Dawson 2013). Analytics also support regulation and redesign by highlighting design elements that need revision before reuse (Hernandez-Leo et al. 2014). During delivery, Learning Analytics can identify student behaviours, allowing lecturers to intervene if these differ from expected outcomes. For example, by sending reminders, offering tutorials, or personalising activities. Such information supports targeted course designs for subgroups of students such as underperforming students, students at-risk, slow learners, or differently abled students (McKenney and Mor 2015; Hansen 2016; Rienties and Toetenel 2016).

At the same time, the need to link analytics approaches and outputs to educational contexts has been widely recognized (Rienties, Toetenel, and Bryan 2015; Gasevic, Dawson, and Siemens 2015). To optimize Learning Analytics for improved student performance and adaptive analysis, the underlying pedagogical context provided by Learning Design must be integrated. Pedagogical plans and objectives from Learning Designs can be appraised against Learning Analytics outputs (Rienties, Toetenel, and Bryan 2015). Without such contextual interpretation, the potential of Learning Analytics is limited (Mangaroska and Giannakos 2017). Learning Design provides this framework, ensuring meaningful analysis of student behavioral data and accurate pedagogical recommendations (Looney and Siemens 2011; Lockyer, Heathcote and Dawson 2013).

Therefore, to enhance higher education, Learning Design and Learning Analytics must be integrated into a coherent cycle. As Ifenthaler, Gibson and Dobozy (2017) observed, one of the next frontiers in Blended Learning research is the synergistic relationship between Learning Design and Learning Analytics.

Different definitions for Learning Analytics can also be found in the literature. The most comprehensive and most cited definitions are given in **Table 2.**

Table 2: Learning Analytics Definitions

Author(s)	Definition
Looney and Siemens (2011)	*"Learning Analytics is the use of intelligent data, learner-produce data, and analysis models to discover patterns and connections within that data, and to predict and advise on learning."*
LAK11 website (2011)	*"Learning Analytics is the measurement, collection, analysis and reporting of data about learners and their contexts, for the purpose of understanding and optimizing learning and the environments in which it occurs."*
Johnson et al. (2011)	*"Refers to the interpretation of a wide range of data produced by and gathered on behalf of students in order to assess academic progress, predict future performance, and spot potential issues."*
Jisc Program (2017)	*"Meeting the challenge of using data and analytics to support students; improving satisfaction, retention and graduation rates."*

The use of Learning Management Systems (LMS), which can store and visually represent student information, enabled the active development of the Learning Analytics field. Learning Analytics apply Data Mining and Machine Learning techniques to identify hidden patterns in LMS data and use information visualization techniques to generate summative, real-time, and predictive feedback to both students and lecturers. For students, Learning Analytics tools provide opportunities for self-evaluation and comparison with peers. For lecturers, these tools provide evidences about students' performance, engagement, potential risks, and the effectiveness of teaching methods. For institutions, analytics support broader decision-making, including student recruitment, curriculum planning, and financial policies (Campbell and Oblinger 2007).

3. Research Methodology

3.1 Purpose of the Systematic Literature Review and Research Questions

The aim of this systematic literature review is to examine and synthesize existing empirical research on the integration of Learning Analytics and Learning Design in higher education. Furthermore, the review seeks to identify the key benefits, challenges, and opportunities arising from this integration. In order to clearly capture the contributions of the Learning Analytics and Learning Design synergy, the following research questions were developed:

RQ1. *What benefits emerge from the integration of Learning Analytics and Learning Design in higher education?*

RQ2. *What opportunities arise from the implementation of synergy between Learning Design and Learning Analytics to enhance teaching and learning in higher education?*

3.2 The Originality of the Literature Review

A search query was formulated with a set of keywords focusing on the targeted research topic to find similar reviews already conducted in this domain (Webster and Watson 2002). Databases such as IEEE Xplore, ACM

Digital Library, Scopus, and Google Scholar were selected for the search. The search query yields four similar literature reviews in the domain, and all reviews were carefully analyzed.

The first research paper, entitled "A review of ten years of implementation and research in aligning learning design with learning analytics at the Open University UK" by Rienties et al. (2017), presented the lessons learned from eight research works conducted on applying Learning Analytics to understand the impact of Learning Designs on student performance, behavior, and satisfaction. However, this review only considered the research works conducted at the Open University UK.

The second research paper, entitled "Learning analytics for learning design: A systematic literature review of analytics-driven design to enhance learning" by Mangaroska and Giannakos (2018), considered how Learning Analytics can support Learning Design. They aimed to investigate the current status of Learning Analytics for Learning Design and classified the Learning Analytics indicators that have been used to inform Learning Design decisions.

The third research paper, entitled "Learning design and learning analytics in mobile and ubiquitous learning: A systematic review" by Pishtari et al. (2020), presented a systematic literature review of Learning Analytics and Learning Design in mobile and ubiquitous learning (M/U-learning) by reviewing 54 papers published in the domain. This literature review was conducted to investigate how M/U-learning, Learning Design, and Learning Analytics are related.

The fourth research paper, entitled "Connecting the dots—A literature review on learning analytics indicators from a learning design perspective" by Ahmad et al. (2022), investigated the role of Learning Design in Learning Analytics through a systematic literature review by analyzing 161 research papers. The aim of this paper is to create a reference framework that connects Learning Analytics and Learning Design and to identify and analyze the Learning Analytics indicators and metrics used over the past decade.

The aim and research questions of our systematic literature review are different from the above literature review papers. To the best of our knowledge, at the time of writing this paper, there were no journal papers or conference publications directly related to our research aim and the research questions.

3.3 Research Process

This study used the eight-step paradigm suggested by Okoli and Schabram (2010) to conduct a systematic literature review. These steps are: (1) Identify the purpose; (2) draft the protocol; (3) apply practical screen; (4) search for literature; (5) extract data; (6) appraise quality; (7) synthesize studies; (8) write the review. All the eight steps are essential for creating a scientifically rigorous systematic literature review (Okoli 2015). The activities undertaken at each of these steps are described below, and **Figure** 1 illustrates the overall systematic literature review process.

Step 1: As mentioned in section 3.1, the purpose of the systematic literature review is to identify empirical evidence that demonstrates the benefits and opportunities that might arise as a result of implementing the synergy between Learning Analytics and Learning Design to enhance higher education.

Step 2: Research protocols were drafted to conduct the literature review, including the following steps.

- Selection of the international databases: IEEE Xplore, ACM Digital Library, Scopus, and Google Scholar.
- The search queries were formulated using the keywords focusing on the targeted research topic. Three main search terms, *'Learning Analytics,' 'Learning Design,'* and *'Higher Education,'* were used.

All search queries contained the above three main terms in combination with the following terms in research paper titles.

('synergy' OR 'alliance' OR 'collaboration' OR 'connect') AND ('education' OR 'teach' OR 'learn', OR 'study success' OR 'retention' OR 'course completion')

- Annotating digital paper copies and linear notes were used as note-taking techniques.

Step 3: The practical screening was applied based on the previously defined inclusion criteria in the research protocol. The literature inclusion criteria of the study are listed below.

- Research studies were positioned in the higher educational context.
- Research papers were either peer-reviewed journal articles or conference papers.
- Presented either qualitative or quantitative analyses and findings.

- Published after 2011.
- Published in the English language.
- An abstract was available.

Step 4: A literature search was conducted based on the keywords mentioned in the protocol. Search queries returned N = 157 publications.

Step 5 & Step 6: These steps were executed in parallel, which includes the following sub-steps.

- An initial search identified N = 157 research papers. After removing irrelevant papers and duplicates, N = 67 papers were discarded, resulting in N = 90 papers retained for abstract-level screening.
- A detailed analysis of the abstracts was performed, focusing on the relevance of key concepts (e.g., Learning Analytics for Learning Design). In this phase, N = 21 papers were excluded, leaving N = 69 papers for full-text review.
- A comprehensive analysis of the full papers was conducted, considering theoretical background, methodology, and experimental evidence. During this stage, N = 14 papers were discarded, resulting in a final set of N = 55 key publications included in the literature review.

Step 7: Relevant information was extracted from the selected research papers by reading the abstract, methodology, experimental results, and pedagogical approaches. Extracted information was coded, summarized, and organized based on the research aim, research questions, research design, and contribution to the domain.

Step 8: Key findings were documented by writing this research paper.

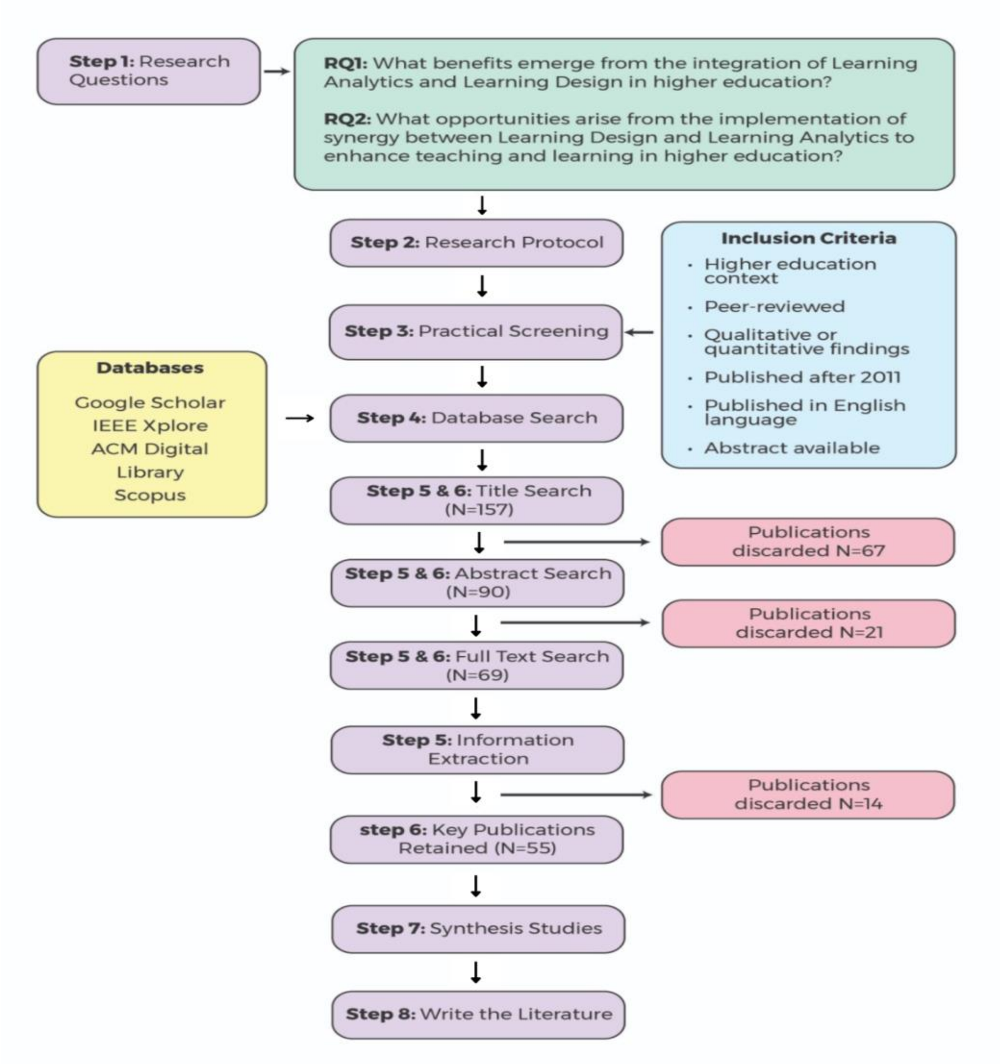


Figure 1: Flow chart of the research process

4. Results

4.1 RQ1: What Benefits Emerge from the Integration of Learning Analytics and Learning Design in Higher Education?

The benefits of using the synergy between the Learning Design and Learning Analytics are associated with three levels: micro-level (students), meso-level (lecturers), and macro-level (higher educational institutions).

Micro Level - Students

According to Rienties and Toetenel (2016), learning is not always an enjoyable experience. Making mistakes and receiving poor grades can make learning challenging at times. As a result, getting positive feedback and assistance is crucial for ongoing learning (Rienties and Toetenel 2016). Learning Analytics visualizations are crucial for boosting students' self-esteem and motivation.

Learning Analytics Dashboard displays facilitate the ongoing enhancement of engagement, performance and satisfaction. While visualizing assignment grades on a line chart encourages students to keep improving, visualizing their engagement with the materials also improves their drive. The combination of Learning Analytics and Learning Design can boost student learning engagement, performance, learning success, career objectives, and retention, according to previous studies (Rienties, Toetenel, and Bryan 2015; Rienties and Toetenel 2016).

According to Ferguson (2012), Mangaroska and Giannakos (2018), the data produced by Learning Analytics tools can boost students' motivation, satisfaction, and confidence. According to Siemens et al. (2011), Mangaroska and Giannakos (2018), students can easily monitor their current progress, identify the subject areas in which they are performing poorly, calculate the time needed for each learning task, and create a customized learning environment based on analytical visualizations. A summary of the important research activities conducted to facilitate students is given in **Table 3.**

Table 3: Summary of research activities conducted to facilitate students.

Author(s)	Year	Aim	LA Methods/ Research Design	Contribution(s)
Rienties, Toetenel and Bryan	2015	How does Learning Design affect performance?	Cluster analysis. Correlation analysis.	Learning Designs strongly influence student engagement and performance.
Atkinson	2015	Empower the learner with their own ability to make adjustments to their self-learning.	Identify student engagement. Visually present using pie charts.	SOLE Model—empowers the Research student to adjust their self-learning environments by using Learning Analytics tools. SOLE toolkit.
Toetenel and Rienties	2016	How do Learning Designs affect LMS behavior, satisfaction, and performance in blended and online environments?	Multiple regression models.	Importance of Learning Designs in predicting and understanding the performance of students.
Toetenel and Rienties	2016	The configuration of Learning Designs and its effect on students' performance.	Cluster analysis. Correlation analysis.	Importance of Learning Designs in LMS behavior and performance.

Author(s)	Year	Aim	LA Methods/ Research Design	Contribution(s)
Wise	2016	How are analytics taken up and used as part of the teaching and learning process?	Focus on students as analytics users.	Student Tuning Model. Align Design framework - a continual cycle model in which students plan, monitor, and change their learning activities as they engage in the learning activities with Learning Analytics.
Schumacher and Ifenthaler	2016	Investigate features students expect from Learning Analytics.	Qualitative exploratory study.	Features students expect are self-assessments, personalized analyses, adaptive recommendations, and support to plan and organize their learning process.
Ifenthaler, Gibson and Dobozy	2017	Demonstrate how analysis could inform Learning Design of the self-guided digital learning experience.	Network graph analysis. Navigation pattern analysis.	Dashboard visualizations provide a self-guided digital learning experience.
Nguyen, Rienties and Toetenel	2017	How was Learning Designs configured longitudinally across modules & what was its effect on LMS behavior?	Visualization Network Analysis. The fixed-effect regression model.	Learning Designs identified the variability in student online activities.
Nguyen, Rienties and Toetenel	2017	How were Learning Designs configured at the activity level, and what media was used to deliver them?	Visualization. Network analysis.	Find out how learning activities interact with each other across modules. Assimilative activities accounted for the majority of study time.
Mavrikis and Karkalas	2017	Increase awareness regarding the use of educational e-books.	Dashboard Visualization.	Reflective Designer Analytics Platform (RDAP) to increase the awareness of the use of e-books.
Charleer et al.	2018	Facilitate communication between study advisers and students by visualizing grade data.	Dashboard Visualization.	LISSA dashboard.
Antonette, Simon and Simon	2019	Develop scalable Learning Analytics applications that can cater to a large number of students.	Design and Creation.	CLAD model for Learning Analytics to align with pedagogical contexts. AcaWriter tool to improve the academic writing of Students.
Wang and Han	2020	Provide process-oriented feedback to students to enhance learning effectiveness.	Dashboard Visualization.	iTutor-Learning Analytics Dashboard.

Author(s)	Year	Aim	LA Methods/ Research Design	Contribution(s)
Vujovic et al.	2020	How the shapes of shared tables affect the learners' behaviour when collaborating in terms of patterns of participation.	Multimodal learning analytics. Quantitative and qualitative analyses.	Use of round tables (vs rectangular tables) leads to higher levels of on-task participation in the case of elementary school students.
Fan et al.	2021	Identify learning tactics and their links with Learning Design elements.	Cluster analysis, process mining technique, and an epistemic network analysis.	Detected four learning tactics (Search oriented, content-and-assessment-oriented, content-oriented, and assessment-oriented) that were used by MOOC learners.
Banihashem et al.	2021	Investigate the effect of the Constructivist Learning Design and Learning Analytics (CLDLA) Model on learners' engagement and self-regulation	Experimental study.	The CLDLA model has a positive impact on learners' engagement and self-regulation.
Duan, Wang and Rouamba	2022	Implement a Learning Analytics Dashboard to generate actionable feedback for students to advance their self-regulated learning skills and improve their grades.	Dashboard Visualization.	Learning Analytics Dashboard.
Jayashanka et al.	2022	Improve motivation, engagement, and grades of students.	Dashboard Visualization.	TELA-System (a Moodle plugin).
Ochukut et al.	2024	Alignment of Learning Design with Learning Analytics in Moodle-based blended learning.	Case study: activity log analysis.	Demonstrated how aligning Learning Designs with Learning Analytics in blended learning improved engagement and completion rates.
Possaghi et al.	2025	Integrate multi-modal Learning Analytics Dashboard in K-12 education.	User-centered design. Multi-modal data analytics.	Developed an LAD for open-ended activities in K-12 settings.

Author(s)	Year	Aim	LA Methods/ Research Design	Contribution(s)
Echeverria et al.	2025	How students from high and low-performing groups reflect individually on their collaborative performance while using a Learning Analytics Dashboard (LAD).	Post-hoc semi-structured individual interview. Bain's 5R reflection framework. Epistemic Network Analysis (ENA).	Use of LADs to provoke reflection must be accompanied by scaffolding strategies. Researchers studying student reflection with Learning Analytics Dashboards should examine not only what and how often students reflect, but also the qualitative depth and stages of their reflection.
Jin et al.	2025	Investigates the role of GenAI literacy in learner interactions with conventional versus scaffolding chatbot-assisted LADs.	2x2 mixed-method experiment. Comparative analysis.	Highlighted the importance of considering learners' GenAI literacy when integrating GenAI chatbots in LADs and educational technologies.
Marques, Hernández-Leo and Castillo	2025	How factors of the educational setting and student performance interplay with student satisfaction with the Learning Design.	Student satisfaction survey. Reliability Assessment Score. Institutional analytics.	Learning Design aspects strongly correlate with students' holistic perception of a course.

Meso level - Lecturers

Feedback from students is typically gathered after learning activities are finished. This makes it impossible for lecturers to intervene in students' learning in real time (Nguyen, Rienties, and Toetenel 2017). The reflection stage of the teaching process is frequently restricted to intuition gleaned from self-reports, course assessments, and evaluations. Real-time interventions may be further limited by response and selection bias in these feedback processes (Nguyen, Rienties, and Toetenel 2017).

By generating feedback that is rich in information, data gathered by Learning Management Systems can empower lecturers and speed up the teaching process. This enables lecturers to make real-time interventions and assess their methods at various granularities (Nguyen, Rienties, and Toetenel 2017). By allowing lecturers to assess the efficacy of Learning Designs and ascertain whether anticipated learning outcomes are attained, the synergy between Learning Analytics and Learning Design facilitates this process. Additionally, these statistics make it easier to identify Learning Design components that need to be revised before being used again.

Learning Analytics data, such as the frequency of subject access or the amount of time spent on learning activities, offer real-time insights into how students respond to a Learning Design (Nguyen, Rienties, and Toetenel 2017). These behavioral indicators enable lecturers to develop tailored interventions, modify lesson plans in response to student trends, spot underachievers early, and offer focused assistance (Law et al. 2017; Hansen 2016; Rienties and Toetenel 2016). To address topics that are difficult to understand, lecturers may also give extra lectures, tutorials, or educational materials.

Learning Design and Learning Analytics work together to promote ongoing enhancements to the quality of instruction. Using analytical visualizations, lecturers can assess teaching methods, engage with students more successfully, and improve autonomy and decision-making. A summarization of the important research activities conducted to facilitate the lecturers is given in **Table 4.**

Table 4: Summary of research activities conducted to facilitate lecturers

Author(s)	Year	Aim	LA Methods/ Research Design	Contribution(s)
Jonathan et al.	2012	Investigate feedback for teaching and learning using analytics via Moodle.	Analyzed: Engagement Involvement Interaction Influence	Implementation of feedback display.
Lockyer, Heathcote and Dawson	2013	How Learning Designs might provide the framework for interpreting Learning Analytics results?	Checkpoint analysis. Process analysis.	Evaluation of learning within a specific pedagogical design. Visualizations provide lead indicators of student engagement and how they progress through the planned learning sequence.
Emin-Martinez et al.	2014	Help teachers to align both the improvement of their practices and the orchestration of their classrooms.	Combines Learning Design and Analytics to improve the adoption and assessment of learning tasks.	Teacher-led design inquiry of learning" as a new model of educational practice and professional development.
Kennedy et al.	2014	Investigate and develop ways in which Learning Analytics can be harnessed by teaching in higher education.	Access new forms of empirical data to teachers.	Conversational Framework - highlighted the interaction, dialogue, and feedback between teachers and students are critical to students' learning process and outcomes. Explore teachers' views on how Learning Analytics might help them address known difficulties. Web-based analytics tool to support teachers.
Rodríguez-Triana et al.	2015	Proposal for alignment between Learning Designs and Learning Analytics to support teachers in designing scenarios.	Scripting. Monitoring.	Exploits the synergies between Learning Designs and Learning Analytics in Computer-Supported Collaborative Learning using monitoring aware design process and script aware monitoring process.
McKenney and Mor	2015	How teachers are supported in the synergistic processes integral to the educational design?	Evaluation of a tool and results of the retrospective analysis.	CASCADE-SEA: A computer-based support tool which provides reflective analysis on Learning Designs.

Author(s)	Year	Aim	LA Methods/ Research Design	Contribution(s)
Hernández-Leo and Pardo	2016	Empower the learner with their own ability to make adjustments to their personal learning.	Articulation of multiple authoring tools.	ILDE community platform to support teachers to design learning activities using multiple authoring tools.
Toetenel and Rienties	2016	The configuration of LD and its effect on students' performance.	Visualization. Correlation analysis.	Seven types of Learning Designs analyzed through Learning Analytics methods. They found out the high focus given for the assimilative and assessment learning activities. Lower usage of student-active activities.
Toetenel and Rienties	2016	Whether collaborative, networked approach changed how educators design courses.	Comparison of 148 prior and post Learning Design initiatives.	By visualizing design upfront, educators focused less on traditional teaching patterns.
Bakharia et al.	2016	Inquiry-based evaluation of Learning Designs.	Temporal Analysis. Tool-specific Analysis. Cohort Analysis. Comparative Analysis	Conceptual Framework—Presents more meaningful data to teachers to evaluate their Learning Designs and to transform Learning Designs into a teacher-led, inquiry-based practice. Loop Tool—the reference implementation of the conceptual framework.
Gunn et al.	2016	Professional development of teachers.	Outlines a professional development initiative.	Make Learning Analytics practice accessible to teachers.
Kitto et al.	2016	Proposes direct solutions for helping people to imagine how Learning Anaytics might be used in a more nuanced manner.	Social network analysis. Content analysis.	CLA Toolkit - uses xAPI to unify the description of data gathered from various media. Three Learning Design patterns to support Learning Analytics protocols.
Nguyen, Rienties and Toetenel	2017	How was Learning Designs configured longitudinally across modules & what was its effect on LMS behavior?	Visualization Network Analysis. Fixed-effect regression model.	Learning Designs identified variability in student online activities. Learning Designs able to explain up to 60% of the variability in student online activities.

Author(s)	Year	Aim	LA Methods/ Research Design	Contribution(s)
Law et al.	2017	Support inter-professional collaboration among learning designers and Learning Analytics communities.	Design and creation.	Learning Design Studio (LDSHE) to facilitate interprofessional collaboration among learning designers and Learning Analytics communities.
Nguyen, Rienties and Toetenel	2017	How were Learning Designs configured at the activity level, and what media was used to deliver them?	Visualization. Network Analysis.	Assimilative activities accounted for majority of study time. Lecturers more likely to use reading materials to convey information.
Marcel et al.	2017	Improve the Learning Design of a course during run-time.	Design Science Research.	Learning Analytics Dashboard.
Inventado and Scupelli	2017	Encourage collaboration among existing communities of stakeholders.	Four objectives for online learning collaboration.	Online learning collaborator framework to encourage collaboration among stakeholders.
Mavrikis and Karkalas	2017	Increase awareness regarding the use of educational e-books.	Dashboard Visualization.	Reflective Designer Analytics Platform (RDAP) - developed for lecturers and designers to support the creation of interactive e-books for learning.
Eradze, Rodríguez-Triana and Laanpere	2017	Introducing observational data into Learning Analytics datasets to provide a more holistic view of the teaching and learning process.	Research-based design process.	Reference Model. Observata—lesson observational tool.
Davinia Hernández-Leo et al.	2018	Propose a framework to support informed decision-making in Learning Design.	Design a framework with three layers of Data Analytics: —Learning Analytics, Design Analytics, and Community Analytics.	AL4LD framework.
Xing et al.	2019	Employ Learning Analytics to build performance prediction models to help struggling students.	Radial Basis Function-based Support Vector Machines and the tree classification method.	Develop a model to identify struggling students and provide actionable insights for teachers to provide personalized and timely feedback to students.

Author(s)	Year	Aim	LA Methods/ Research Design	Contribution(s)
Kaliisa, Kluge and Mørch	2020	Investigate how Learning Analytics can inform Learning Designs in Blended Learning environments.	Social Network Analysis. Text Network Analysis.	Analyzing different levels of analytics could provide important information about student online learning processes, which can be used as a reflective resource by teachers to make informed Learning Designs decisions.
Wang and Han	2020	Provide process-oriented feedback to students to enhance learning effectiveness.	Dashboard Visualizations.	iTutor-Learning Analytics Dashboard.
Nguyen et al.	2021	Develop, demonstrate, and evaluate a set of design principles for Information Systems (IS) that utilize Learning Analytics to support learning and teaching in higher education.	Design Science Research Methodology.	Established a foundation for further development and implementation of Learning Analytics Information Systems (LAIS) in higher education.
Mohseni, Martius and Masiello	2021	Explore the design of a Learning Analytics Dashboard with the use of interactive visualizations and Machine Learning.	Data Visualizations.	SAVis tool.
Stoyanov and Paul	2023	Aims to identify the most relevant concepts at the intersection of Learning Designs and Learning Analytics.	Critical interpretive synthesis. Qualitative content analysis. Text analytics.	Identified two topics rarely explicitly discussed in the literature: 'evidence-informed instructional design approaches' and 'design-based research.'
Pelizzari, Sala and Tassalini	2024	Pilot implementation of Learning Analytics in Higher Education for teaching improvement.	Case study. Data infrastructure. Role-based dashboards.	Demonstrated Learning Analytics positive impact on teaching practices, early risk detection, and institutional decision support.

Author(s)	Year	Aim	LA Methods/ Research Design	Contribution(s)
Liu et al.	2024	Assist K-12 teachers in assessing students' collaborative problem solving skills in an educational game.	Survey. Partial Least Squares Structural Equation Modelling (PLS-SEM).	A Learning Analytics Dashboard was implemented to assist K-12 teachers. Makes theoretical, methodological, and practical contributions to technology integration in Learning Analytics Dashboard implementation.
Possaghi et al.	2025	Integrate multi-modal Learning Analytics Dashboard in K-12 education.	User-centered design. Multi-modal data analytics.	Developed an Learning Analytics Dashboard for open-ended activities in K-12 settings.
Alenezi and Alenezi	2025	Applies data analysis to the improvement of online course design.	Descriptive analytical approach employing mixed method. Survey.	Learning Analytics possess substantial potential to transform online course design and foster better student outcomes.

Macro Level – Higher Educational Institutes

Curriculum creation and the launch of new academic programs can benefit from the data produced by the synergy between Learning Analytics and Learning Design. Furthermore, learner cohort analysis, lower dropout rates, improved student retention, and overall academic achievement can all be achieved with the use of such data (Ifenthaler and Widanapathirana 2014). Applications at the institutional level have received very little attention, despite the fact that the majority of current research focuses on the advantages at the student and lecturer levels. This shows a glaring research gap and emphasizes the need for further empirical studies that look at how higher education institutions might strategically use Learning Analytics in line with Learning Design to guide long-term planning and decision-making. The observed benefits may be methodically traced across three hierarchical levels because the main stakeholders mentioned in the literature are students, lecturers, and higher education institutions (Ifenthaler and Widanapathirana 2014). This classification makes it easier to find research gaps and allows for an organized assessment of the benefits of this synergy. The benefits found in the literature are summarized in Table 5 by mapping them across three hierarchical stakeholder levels: students (micro-level), lecturers (meso-level), and higher education institutions (macro-level).

Table 5: Benefits of synergy between Learning Analytics and Learning Design for different stakeholders

Stakeholder	Benefits
Student	Check current performance level. Indicate key materials to learn. Facilitate the creation of a personalized learning environment. Indicate subject areas need to improve. Visualize the required amount of study time. Track the progress towards learning goals. Increase engagement. Improve grades. Improve motivation and self-confidence. Optimize learning activities.

Stakeholder	Benefits
Lecturer	Increase the quality of teaching. Identify students at-risk of failure/underperforming students. Adjust the learning materials to the needs of the learners. Create meaningful interventions. Modify the Learning Designs to meet cohorts' needs. Compare Learning Designs. Identify Learning Designs that need to revise. Evaluate Learning Materials. Plan for future interventions.
Higher Educational Institutes	Adjustments of the curriculum to the needs of learners. Increase the quality of the curriculum. Optimize resource allocation. Model retention rates. Support to take financial decisions. Minimize the dropout rates.

A methodical evaluation of how Learning Analytics and Learning Design work together to serve various players in the higher education is made possible by this organized classification. Fewer empirically confirmed benefits are documented at the institutional level, despite a significant number of benefits being reported at the student and lecturer levels. This gap emphasizes the narrow focus of earlier research on macro-level outcomes including strategic decision-making, curriculum-wide optimization, and policy formation rather than suggesting a lack of potential benefit for higher education institutions. As a result, **Table 5** serves as an analytical tool to discover unexplored regions and direct future research on Learning Analytics-informed Learning Design applications at the institutional level, in addition to summarizing recognized benefits.

4.2 RQ2: What Opportunities Arise from the Implementation of Synergy Between Learning Design and Learning Analytics to Enhance Teaching and Learning in Higher Education?

The systematic literature review reveals four primary opportunity areas emerging from the integration of Learning Analytics and Learning Design. These opportunities signify a shift from intuition-based teaching to evidence-informed, adaptive, and ethically governed educational practices. The following section provides a detailed examination of each of these four opportunity areas.

- Making Learning Designs more reliable for lecturers and students

The participatory involvement of all stakeholders in the learning community is important in planning suitable practices to implement design models and the tools to support them (Emin- Martinez et al. 2014; Matcha et al. 2019). There is a growing need to establish an online repository of Learning Designs that can be accessible by a wide educational community. Also, it is important to implement a participatory culture in the field of Learning Design (Emin-Martinez et al. 2014; Persico and Pozzi 2015). Learning Design helps lecturers to describe, communicate, and share their designs with the Learning Design community. Good teaching practices from one educational context can be captured and reused in another context (Lockyer, Heathcote, and Dawson 2013). By sharing Learning Designs within the Learning Design community, the lecturers can comprehend whether their particular design or model leads to an effective learning experience for the learner. At the same time, lecturers should be encouraged to review peer Learning Designs and provide feedback.

In current learning environments, students' feedback is collected after studying the course module. Rienties et al. (2017) highlighted the importance of getting the involvement of students' participation in the Learning Design activities. Open University's Learning and Teaching Innovation curriculum design student panel is a novel approach to get the students' contribution to the model development using focus groups (Rienties et al. 2017).

Recently, design patterns have been suggested as a construct to mediate between Learning Design and Learning Analytics (Inventado and Scupelli 2015). The Learning Analytics community needs to create a pattern repository to support lecturers as a source of inspiration when creating new course content (Kitto et al. 2016). Design patterns should support the transfer of the currently available effective Learning Designs to other learning contexts without reinventing the wheel (Antonette, Simon, and Simon 2019). Research work that has been conducted in terms of the practices, tools, and representations to evaluate the effects of Learning Designs is limited. Learning Design needs to incorporate built-in evaluation methods to analyze whether the expected outcomes were achieved. Lecturers should be able to generate designs that are compatible with emerging Learning Analytical technologies and tools (Atkinson 2015). Lecturers need to design units of learning that can be deconstructed and rebuilt in meaningful ways to enable the Learning Analytic algorithms to function optimally (Atkinson 2015).

- Design and implementation of smart Learning Analytics tools

The design and development of smart Learning Analytics tools that can provide summative, real-time, and predictive feedback to students, lecturers, and educational institutes is another opportunity in this domain (Matcha et al., 2019; Valle et al., 2021). These tools and frameworks play a vital role when creating a synergy between Learning Designs and Learning Analytics, which in turn provides benefits to enhance higher education. As stated in **Table 3** and **Table 4,** researchers developed different technologies (tools, frameworks, and models) to enhance the learning environments through the synergy between Learning Analytics and Learning Design. However, most of those tools are in the development and experimental stages. The opportunities that arise related to the design and implementation of smart Learning Analytic tools are mentioned below.

The first opportunity is about the data and the environment. There is a growing need to develop methods to work with a wide range of datasets, shift towards more challenging datasets, and combine datasets which include mobile data, biometric data, and mood data to improve learning environments. Benefits of Learning Analytic tools can improve by using the physical data which are not directly linked with educational data, such as the student's current emotional state, self-confidence, demographic, and socio-cultural data (Rienties, Toetenel and Bryan, 2015; McKenney and Mor, 2015). The implementation of tools and analytic methods to handle big data and deliver real-time meaningful results is also important.

The second opportunity is about defining the right objectives/indicators/metrics triple before the implementation of Learning Analytics tools. Focusing on objectives concerning the learners' perspectives is essential to the development of tools related to learners' needs, rather than the needs of institutions (Ferguson, 2012; Williamson and Kizilcec, 2022). These perspectives can extend the criteria for learning success beyond grades and persistence to include motivation, confidence, satisfaction, and meeting career goals (Williamson and Kizilcec, 2022).

The third opportunity is to develop tools that can easily be used for analysis and visualizations without having extensive knowledge of the techniques underlying these tools. It is important to consider the requirements of both students and lecturers when designing and implementing analytic tools. Both lecturers and students can easily understand simple dashboard visualizations (Valle et al., 2021; Matcha et al., 2019). It is really important to keep in mind that students are different and unique when designing and implementing Learning Analytics tools. Students can be empowered to use his/her own ability to refine their learning environment by using Learning Analytics tools (Atkinson, 2015). It is important to improve the usability of analytic tools and create user manuals and guidelines to build a good interaction between the users (lecturers and students) and the analytic tools.

- Professional Development of Lecturers

The need to improve the use of "learning intelligence" in lecturers is becoming urgent due to an interdisciplinary space between computer researchers and lecturers (Melero et al. 2015; Celik and Magoulas 2016). Lecturers are curious about experimenting with new methods to integrate Learning Analytic tools with Learning Designs. For this, they have to improve their programming skills needed to develop such technologies (Celik and Magoulas 2016). They frequently underrate the potential of data science, since they remain unaware of the different methods by which educational datasets can be analyzed and visualized (Celik and Magoulas 2016). Therefore, it is essential to increase data literacy among lecturers (Melero et al., 2015). Professional development, training, incentives, and rewards are critical for the successful adoption of Learning Analytics tools (Cathy et al., 2016). Cathy et al. (2016) presented Learning Analytics professional development scenarios, and they highlighted the importance of expertise to support lecturers to develop their data literacy skills (Cathy et al. 2016).

- Development of a clear set of ethical guidelines

Another opportunity that emerges when implementing the synergy between Learning Design and Learning Analytics is ethical, legal, and risk considerations. The development and application of a clear set of ethical guidelines are important. Under this, data accuracy, how to respect privacy, maintaining anonymity, preventing data misuse, protecting confidential user information, data ownership, data preservation, sharing data with outside parties, and proper training for educational practitioners regarding the handling of data should be taken into consideration.

Figure 2 provides a graphical summary of the key findings from this Systematic Literature Review, highlighting the benefits and opportunities resulting from the integration of Learning Analytics and Learning Design.

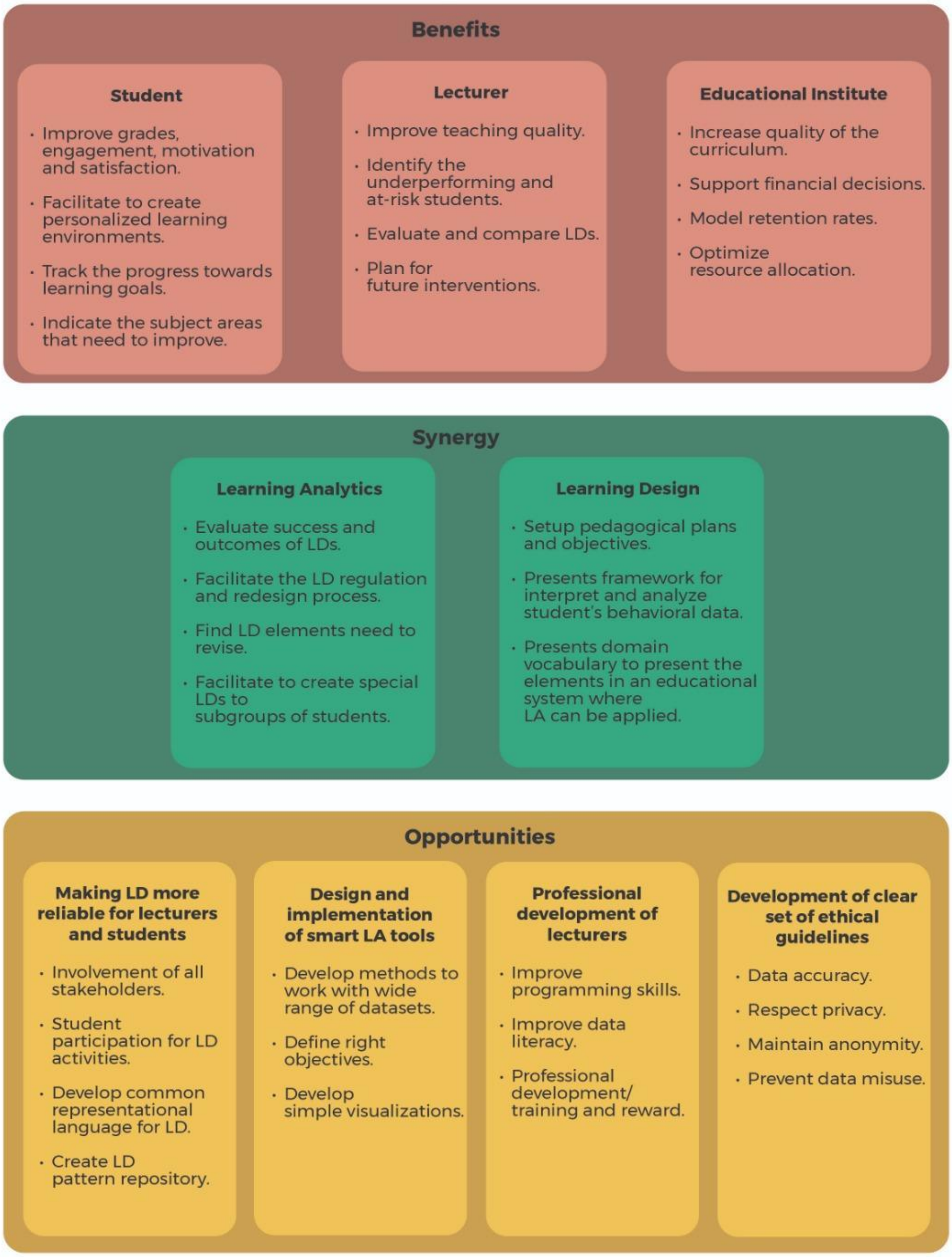


Figure 2: Benefits and Opportunities Arise as a Result of implementing the Synergy Between Learning Analytics and Learning Design

5. Conclusion

By incorporating computer technologies into teaching methods, universities are increasingly encouraging their faculty to innovate. This comprehensive literature review investigated the relationship between Learning Analytics, the use of data science in educational settings, and Learning Design, the framework for creating interactive learning experiences. Our analysis demonstrates a synergistic relationship: Learning Analytics produces the empirical evidence needed to assess and improve Learning Designs, while Learning Designs supply the instructional goals and semantic framework.

This systematic literature review found important advantages for students, lecturers, and higher-educational institutions based on a thorough analysis of 55 systematically selected research papers. Self-Regulated Learning models (like SOLE toolkit), decision-making frameworks (like AL4LD), and engagement dashboards (like TELA-System) are important contributions. Additionally, putting this synergy into practice offers four transformative opportunities: Making Learning Designs more reliable for lecturers and students, Design and implementation of smart Learning Analytics tools, Professional Development of Lecturers, and Development of a clear set of ethical guidelines.

Notwithstanding these developments, the analysis identifies a number of crucial topics that require further research. There is an urgent need to implement intelligent and interactive Learning Analytics Dashboards for students that provide real-time feedback and support customized learning pathways, as the majority of existing tools remain lecturer-centric. Additionally, there is still a lack of study on the use of advanced AI, such as Artificial Neural Networks, Deep Learning and Generative AI to enhance learning environments.

Future studies should also use sensing technologies that record affective characteristics, such as student motivation and emotion, in addition to typical clickstream data. Lastly, research addressing the macro-level requirements of higher-educational institutions and national policy formulation is conspicuously lacking, despite the well-documented micro-level benefits of classrooms. The long-term development of analytics-informed Learning Designs in higher education will depend on filling in these gaps and creating explicit ethical standards for data protection.

AI Statement: The authors declare that they used ChatGPT and Grammarly to enhance the spelling, language quality, and overall readability of this manuscript.

Ethical Declaration: This study did not require ethical approval because it is a Systematic Literature Review based exclusively on previously published scholarly sources. The research did not involve the collection of privacy data, nor did it include any interaction with human participants or animals. Consequently, no personal, sensitive, or identifiable data were accessed or analyzed.

References

Agostinho, S., 2006. The use of visual learning design representation to document and communicate teaching ideas. *Proceedings of ASCILITE Conference*, Sydney, NSW, pp.1–9.

Ahmad, A., Schneider, J., Griffiths, D., Biedermann, D., Schiffner, D., Greller, W. and Drachsler, H. (2022) 'Connecting the dots: A literature review on learning analytics indicators from a learning design perspective', Journal of Computer Assisted Learning, 40(6), pp. 2432–2470. doi: 10.1111/jcal.12716.

Aksovaara, S., Määttä, S., Kärkkäinen, T. and Silvennoinen, M. (2024) Enhancing study experience through teacher response: a learning analytics case study of two course implementations. Seminar.net, 20(2).

Alenezi, A. and Alenezi, A. (2025) Data analytics in online course design and student engagement. Revista Eduweb, 19(3), pp. 139–152.

Alhadad, S.S.J. and Thompson, K. (2017) 'Understanding the mediating role of teacher inquiry when connecting learning analytics with design for learning', Interaction Design and Architecture(s) Journal (IxD&A), 33, pp. 54–74.

Antonette, S., Simon, K. and Simon, B.S., 2019. Contextualizable learning analytics design: a generic model and writing analytics evaluations. *Proceedings of the 9th International Learning Analytics and Knowledge Conference (LAK 2019)*, Tempe, Arizona, USA, pp.120–130.

Atkinson, S.P., 2015. Using learning design to unleash the power of learning analytics. *Proceedings of ASCILITE 2015*, Perth, pp.6–10.

Bakharia, A., Corrin, L., de Barba, P., Kennedy, G., Gašević, D., Dawson, S., Williams, S., Mulder, R., Williams, D., Campbell, J. and Lockyer, L., 2016. A conceptual framework linking learning design with learning analytics. *Proceedings of the 6th International Conference on Learning Analytics & Knowledge (LAK 2016)*, Edinburgh, 25–29 April, pp.1–10.

Banihashem, S.K., Farrokhnia, M., Badali, M. and Noroozi, O., 2021. The impacts of constructivist learning design and learning analytics on students' engagement and self-regulation. *Innovations in Education and Teaching International*, 59(4), pp.442–452.

Campbell, J.P. and Oblinger, D.G., 2007. Academic analytics. *EDUCAUSE Review*, 42(4), pp.40–57.

Cathy, G., Jenny, M., Claire, D., Marion, B. and John, M., 2016. The missing link for learning from analytics. *ASCILITE Conference Proceedings*, Adelaide, pp.1–9.
Celik, D. and Magoulas, G., 2016. A review, timeline, and categorization of learning design tools. In: *Advances in Web-Based Learning ICWL 2016*. Springer, pp.3–13.
Chacón, J. and Hernández-Leo, D., 2014. Learning design family tree to back reuse and cooperation. *Proceedings of the 9th International Conference on Networked Learning*, Edinburgh, pp.510–517.
Charleer, S., Moere, A., Klerkx, J., Verbert, K. and Laet, T., 2018. Learning analytics dashboards to support adviser–student dialogue. *IEEE Transactions on Learning Technologies*, 11(4), pp.389–399.
Conole, G., 2013. *Designing for learning in an open world*. New York, NY: Springer.
Conole, G., Dyke, M., Oliver, M. and Seale, J., 2004. Mapping pedagogy and tools for effective learning design. *Computers & Education*, 43, pp.17–33.
Craft, B. and Mor, Y., 2012. Learning design: reflections upon the current landscape. *Research in Learning Technology*, 20, pp.85–94.
Dalziel, J., Conole, G., Wills, S., Walker, S., Bennett, S., Dobozy, E., Cameron, L. and Badilescu-Buga, E., 2016. The Larnaca Declaration on Learning Design. *Journal of Interactive Media in Education*, 2016(1), p.7.
Dobozy, E., 2013. Learning design research: advancing pedagogies in the digital age. *Educational Media International*, 50(1), pp.63–76.
Duan, X., Wang, C. and Rouamba, G., 2022. Designing a learning analytics dashboard to provide students with actionable feedback and evaluating its impacts. *Proceedings of CSEDU 2022 (2)*, pp.117–127.
Echeverria, V., Fernandez Nieto, G., Zhao, L., Palominos, E., Srivastava, N., Gašević, D., Pammer-Schindler, V. & Martinez-Maldonado, R., 2025. A learning analytics dashboard to support students' reflection on collaboration. Journal of Computer Assisted Learning, 41(1), e13088. https://doi.org/10.1111/jcal.13088.
Emin-Martinez, V., Hansen, C., Rodríguez-Triana, M.J., Wasson, B., Mor, Y., Dascălu, M., Ferguson, R. and Pernin, J. (2014) 'Towards teacher-led design inquiry of learning', eLearning Papers, 36, pp. 1–13.
Eradze, M., Rodríguez-Triana, M.J. & Laanpere, M., 2017. Semantically annotated lesson observation data in learning analytics datasets: a reference model. Interaction Design and Architecture(s) Journal (IxD&A), 33, pp. 75–91.
Fan, Y., Matcha, W., Uzir, N.A., Wang, Q. and Gašević, D., 2021. Learning analytics to reveal links between learning design and self-regulated learning. *International Journal of Artificial Intelligence in Education*, 31, pp.980–1021.
Ferguson, R., 2012. Learning analytics: drivers, developments, and challenges. *International Journal of Technology Enhanced Learning*, 4(5–6), pp.304–317.
First International Conference on Learning Analytics and Knowledge, 2011. *LAK 2011*. Available at: https://dl.acm.org/doi/proceedings/10.1145/2090116 (Accessed: 21 July 2024).
Gasevic, D., Dawson, S. and Siemens, G., 2015. Let's not forget: learning analytics are about learning. *TechTrends*, 59(1), pp.64–71.
Gunn, C., Blumenstein, M., Donald, C., McDonald, J. and Milne, J., 2016. The missing link for learning analytics. *Show Me the Learning, Proceedings ASCILITE*, Adelaide, pp.255–260.
Hale, F., 2016. *Learning design: details of the developing Edinburgh Learning Design roadmap (ELDeR) at the University of Edinburgh*. Edinburgh: University of Edinburgh. [online] (Accessed: 21 July 2024).
Hansen, R., 2016. Learning analytics: data for learning design. *Journal of Digital Learning in Teacher Education*, 32(2), pp.48–49.
Hernández-Leo, D. & Pardo, A., 2016. Towards integrated learning design with across-spaces learning analytics: a flipped classroom example. *Proceedings of the First International Workshop on Learning Analytics Across Physical and Digital Spaces co-located with the 6th International Conference on Learning Analytics & Knowledge* (LAK 2016), 25–29 April 2016, Edinburgh, UK. CEUR Workshop Proceedings, Vol. 1601, pp. 74–78.
Hernández-Leo, D., Martinez-Maldonado, R., Pardo, A., Muñoz-Cristóbal, J.A. and Rodríguez-Triana, M.J., 2018. Analytics for learning design: a layered framework and tools. *British Journal of Educational Technology*, 50(1), pp.139–152.
Hernandez-Leo, D., Moren, P., Chacon, J. and Blat, J., 2014. LDShake support team-based learning design. *Computers in Human Behaviour*, 37, pp.402–412.
Ifenthaler, D., 2016. Are higher education institutions prepared for learning analytics? *AECT Conference Proceedings*, pp.1–9.
Ifenthaler, D. and Gibson, D.C., 2018. The dynamics of learning engagement in challenge-based online learning. *Proceedings of the 18th IEEE International Conference on Advanced Learning Technologies*, Mumbai, India, pp.1–10.
Ifenthaler, D., Gibson, D. and Dobozy, E., 2017. The synergistic and dynamic relationship between learning design and learning analytics. *ASCILITE Conference Proceedings*, pp.1–9.
Ifenthaler, D. and Widanapathirana, C., 2014. Development and validation of a learning analytics framework: two case studies using support vector machines. *Technology, Knowledge and Learning*, pp.221–240.
Inventado, P.S. and Scupelli, P. (2015) Towards an open, collaborative repository for online learning system design patterns. eLearning Papers, 42, pp. 14–27.
Inventado, P.S. & Scupelli, P., 2017. *An online learning collaboratory to address multidisciplinary learning challenges at scale*. Interaction Design and Architecture(s) Journal – IxD&A, 33, pp.11–32.
Jayashanka, R., Hettiarachchi, E. and Hewagamage, K.P., 2022. Technology enhanced learning analytics dashboard in higher education. *The Electronic Journal of e-Learning*, 20(2), pp.151–170.

Jin, Y., Yang, K., Yan, L., Echeverría, V., Zhao, L. & Alfredo, R., 2025. Chatting with a learning analytics dashboard: the role of generative AI literacy on learner interaction with conventional and scaffolding chatbots. In: Proceedings of the 15th International Learning Analytics and Knowledge Conference (LAK '25), 3–7 March 2025, Dublin, Ireland.

Jonathan, P., San Diego, Ballard, J., et al., 2012. Do Moodle analytics have a role to play in learning design, feedback, and assessment? *Moodle Research Conference*, 14–15 September, pp.1–10.

Johnson, L., Smith, R., Willis, H., Levine, A. and Haywood, K. (2011) The 2011 Horizon Report. Austin, TX: The New Media Consortium.

Kaliisa, R., Kluge, A. and Mørch, A., 2020. Combining checkpoint and process learning analytics to support learning design decisions in blended learning environments. *Journal of Learning Analytics*, 7(3), pp.1–25.

Kennedy, G., Corrin, L., Lockyer, L., Dawson, S., Williams, D., Mulder, R., Khamis, S. and Copeland, S., 2014. Completing the loop: returning learning analytics to teachers. In: *Rhetoric and Reality: Critical perspectives on educational technology, Proceedings ASCILITE*, Dunedin, pp.436–440.

Kitto, K., Lupton, M., Davis, K. and Waters, Z., 2016. Incorporating student-facing learning analytics into pedagogical practice. *Show Me the Learning, Proceedings ASCILITE*, pp.255–260.

Koper, R., 2006. Current research in learning design. *Journal of Educational Technology & Society*, 9(1), pp.13–22.

Law, N., Li, L., Farias Herrera, F., Chan, A. and Pong, T.C., 2017. A pattern language-based learning design studio for an analytics informed inter-professional design community. *IxD&A*, 33, pp.92–112.

Liu, Y., Hu, X., Ng, J. T. D., Ma, Z. and Lai, X. (2025) Ready or not? Investigating in-service teachers' integration of learning analytics dashboard for assessing students' collaborative problem solving in K–12 classrooms. Education and Information Technologies, 30, pp. 1745–1776.

Lluch Molins, L. and Cano Garcia, E., 2023. How to embed SRL in online learning settings? Design through learning analytics and personalized learning design in Moodle. *Journal of New Approaches in Educational Research*, 12(1), pp.120–138.

Lockyer, L., Heathcote, E. and Dawson, S., 2013. Informing pedagogical action: aligning learning analytics with learning design. *American Behavioral Scientist*, 57(10), pp.1439–1459.

Looney, J. and Siemens, G., 2011. Assessment competency: knowing what you know and learning analytics. *Promethean Thinking Deeper Research Paper*, No.3.

Matcha, W., Uzir, N., Gasevic, D. and Pardo, A. (2019) A systematic review of empirical studies on learning analytics dashboards: a self-regulated learning perspective. IEEE Transactions on Learning Technologies, 13(2), pp. 226–245

Magoulas, G.D. and Celik, D., 2016. Approaches to design for learning. *ICWL 2016, LNCS*, pp.14–19.

Mavrikis, M. & Karkalas, S., 2017. Reflective analytics for interactive e-books. Interaction Design and Architecture(s) Journal – IxD&A, 33, pp.33–53.

Marcel, S., Evelien, L., Wolfgang, G., Peter, S. and Hendrik, D., 2017. Opportunities and challenges in using learning analytics in learning design. *EC-TEL 2017 Proceedings*, pp.209–223.

Marques, F., Hernández-Leo, D. and Castillo, C. (2025) Beyond bias in student satisfaction surveys: exploring the role of grades and satisfaction with the learning design. Journal of New Approaches in Educational Research, 14, Article 9

Mangaroska, K. and Giannakos, M. (2017) 'Learning analytics for learning design: Towards evidence-driven decisions to enhance learning', in Proceedings of the 12th European Conference on Technology Enhanced Learning (EC-TEL 2017). Lecture Notes in Computer Science, vol. 10474. Cham: Springer, pp. 428–433.

Mangaroska, K. and Giannakos, M., 2018. Learning analytics for learning design: a systematic literature review of the analytics-driven design to enhance learning. *IEEE Transactions on Learning Technologies*, 12(4), pp.516–534.

McKenney, S. and Mor, Y. (2015) 'Supporting teachers in data-informed educational design', British Journal of Educational Technology, 46(2), pp. 265–279.

Melero, J., Hernández-Leo, D., Sun, J., Santos, P. and Blat, J. (2015) 'How was the activity? A visualization support for a case of location-based learning design', British Journal of Educational Technology, 46(2), pp. 317–329.

Mohseni, Z., Martins, R.M. & Masiello, I., 2021. SAVis: a learning analytics dashboard with interactive visualization and machine learning. In: Proceedings of the Nordic Learning Analytics (Summer) Institute 2021, CEUR Workshop Proceedings, Vol. 2985, Stockholm, 23 August 2021.

Nguyen, Q., Rienties, B. and Toetenel, L. (2017a) 'Mixing and matching learning design and learning analytics', in Zaphiris, P. and Ioannou, A. (eds.) Learning and Collaboration Technologies. Lecture Notes in Computer Science, vol. 10296. Cham: Springer, pp. 1–15.

Nguyen, Q., Rienties, B. and Toetenel, L. (2017b) 'Unravelling the dynamics of instructional practice: A longitudinal study on learning design and VLE activities', in Proceedings of the 7th International Conference on Learning Analytics and Knowledge (LAK 2017). Vancouver, Canada, pp. 168–177.

Nguyen, A., Tuunanen, T., Gardner, L. & Sheridan, D., 2021. Design principles for learning analytics information systems in higher education. European Journal of Information Systems, 30(5), pp. 541–568. DOI: 10.1080/0960085X.2020.1816144

Nicolae, N., Mihai, D. and Stefan, T. (2019) 'Joining informal learning in online knowledge communities and formal learning in higher education: Instructional design and evaluation of a blended learning seminar with learning analytic support', Interaction Design and Architecture(s) Journal (IxD&A), 43, pp. 110–127.

Ochukut, S.A., Hernedez-Leo, D., Oboko, R.O. and Miriti, E., 2024. Alignment of learning design with learning analytics: case of Moodle blended learning course in higher education. *International Conference on Methodologies and Intelligent Systems for Technology Enhanced Learning*, Springer, pp.250–263.

Okoli, C. (2015) 'A guide to conducting a standalone systematic literature review', Communications of the Association for Information Systems, 37(43), pp. 879–910.

Okoli, C. and Schabram, K. (2010) 'A guide to conducting a systematic literature review of information systems research', Sprouts: Working Papers on Information Systems, 10(26), pp. 1–49.

Pelizzari, F., Sala, C. and Tassalini, E. (2024) Design and implementation of learning analytics in higher education: a pilot case study. In: Proceedings of INTED 2024 (International Technology, Education and Development Conference), Valencia, Spain, 4–6 March 2024. Valencia: IATED, pp. 896–904.

Persico, D. and Pozzi, F. (2015) Informing learning design with learning analytics to improve teacher inquiry. British Journal of Educational Technology, 46(2), pp. 230–248.

Pishtari, G., Rodríguez-Triana, M.J., Sarmiento-Márquez, E.M., Pérez-Sanagustín, M., Ruiz-Calleja, A., Santos, P., Prieto, L.P., Serrano-Iglesias, S. and Väljataga, T. (2020) 'Learning design and learning analytics in mobile and ubiquitous learning: A systematic review', British Journal of Educational Technology, 51(4), pp. 1078–1100. doi: 10.1111/bjet.12944.

Possaghi, I., Vesin, B., Zhang, F., Sharma, K., Knudsen, C., Bjørkum, H. & Papavlasopoulou, S., 2025. Integrating multi-modal learning analytics dashboard in K-12 education: insights for enhancing orchestration and teacher decision-making. Smart Learning Environments, 12, p.53. https://doi.org/10.1186/s40561-025-00410-4.

Rienties, B. and Toetenel, L. (2016) 'The impact of 151 learning designs on student satisfaction and performance: Social learning (analytics) matters', in Proceedings of the 6th International Conference on Learning Analytics and Knowledge. Edinburgh, UK. New York: ACM, pp. 339–343.

Rienties, B., Toetenel, L. and Bryan, A. (2015) 'Scaling up learning design: Impact of learning design activities on LMS behaviour and performance', in Proceedings of the 5th International Conference on Learning Analytics and Knowledge (LAK '15). New York: ACM, pp. 315–319.

Rienties, B., Quan, N., Wayne, H. and Katharine, R., (2017) A review of ten years of implementation and research in aligning learning design with learning analytics at the Open University UK. *Interaction Design and Architecture(s) Journal (IxD&A)*, 33, pp.134–154.

Rodríguez-Triana, M.J., Martínez-Monés, A., Asensio-Pérez, J.I. & Dimitriadis, Y., 2015. Monitoring-Aware Learning Design Process: Pilot Studies in Authentic CSCL Scenarios. Computing & Informatics, 34(3), pp.615–651.

Schumacher, C. & Ifenthaler, D., 2016. Features students really expect from learning analytics. Proceedings of the 13th International Conference on Cognition and Exploratory Learning in Digital Age (CELDA 2016), Mannheim, Germany. Available at: https://files.eric.ed.gov/fulltext/ED571398.pdf (Accessed: 23 July 2025).

Siemens, G., Gašević, D., Haythornthwaite, C., Dawson, S., Shum, S.B., Ferguson, R., Duval, E., Verbert, K. and Baker, R.S.J.D. (2011) Open learning analytics: An integrated and modularized platform. Milton Keynes: Open University Press.

Stoyanov, S. and Paul, K., 2023. Text analytics for uncovering untapped ideas at the intersection of learning design and learning analytics: critical interpretative synthesis. *Journal of Computer Assisted Learning*, 39(2), pp.1–20.

Toetenel, L. and Rienties, B., 2016a. Analysing 157 learning designs using learning analytic approaches as a means to evaluate the impact of pedagogical decision-making. *British Journal of Educational Technology*, 47(5), pp.981–992.

Toetenel, L. and Rienties, B., 2016b. Learning design–creative design to visualize learning activities. *Open Learning: The Journal of Open, Distance and e-Learning*, 31(3), pp.233–244.

Valle, N., Antonenko, P., Dawson, K. and Huggins-Manley, A.C., 2021. Staying on target: a systematic literature review on learner-facing learning analytics dashboards. *British Educational Research Association*, pp.1–25.

Vujovic, M., Hernández-Leo, D., Tassani, S. and Spikol, D., 2020. Round or rectangular tables for collaborative problem solving? a multimodal learning analytics approach. *Computers & Education*, 144, p.103699.

Wang, D. & Han, H., 2020. Applying learning analytics dashboards based on process-oriented feedback to improve students' learning effectiveness. Journal of Computer Assisted Learning, 37(2), pp.487–499. https://doi.org/10.1111/jcal.12502

Webster, J. and Watson, R.T. (2002) 'Analyzing the past to prepare for the future: Writing a literature review', MIS Quarterly, 26(2), pp. 13–23.

Williamson, K. and Kizilcec, R. (2022) A review of learning analytics dashboard research in higher education: implications for justice, equity, diversity, and inclusion. In: Proceedings of LAK22: 12th International Learning Analytics and Knowledge Conference, pp. 260–270.

Wise, A.F., Vytasek, J.M., Hausknecht, S., Zhao, Y., Liu, Y. and Bowers, A.J. (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), pp. 155–183.

Xing, W., Pei, B., Li, S., Chen, G. & Xie, C., 2019. Using learning analytics to support students' engineering design: the angle of prediction. Interactive Learning Environments, pp. 1–18. DOI: 10.1080/10494820.2019.1680391

From Chalkboards to Smart Classrooms: Faculty Perceptions of IoT Integration in Jordanian Universities

Tahani Abu Jraiban[1], Yousef Sawalha[2], Ghada Mohammad Suleiman Alukool[3], Raed Salem Alsaraereh[4] and Rawan Alkhabayba[5]

[1]Department of Basic Sciences (Humanities and Scientific), Faculty of Arts, Al-Zaytoonah University of Jordan, Jordan

[2]Journalist and Editorial Translator, UAE

[3]Department of Educational Planning and Research / Ministry of Education / Department of Primary Education Teacher, Faculty of Arts, Al-Zaytoonah University of Jordan, Jordan

[4]Department of Primary Education Teacher, Faculty of Arts, Al-Zaytoonah University of Jordan, Jordan

[5]Department of Basic Sciences, Faculty of Arts, Al-Zaytoonah University of Jordan, Jordan

3249@zuj.edu.jo (Corresponding author)

https://doi.org/10.34190/ejel.24.2.4503

Abstract: Digital transformation in higher education has increased interest in faculty adoption of emerging technologies such as the Internet of Things (IoT). This study investigates faculty perceptions of IoT integration in Jordanian private universities, with particular attention to gender and academic rank. Grounded in the Technology Acceptance Model (TAM) and the Unified Theory of Acceptance and Use of Technology (UTAUT), the study examines how key acceptance constructs shape IoT adoption in teaching. A quantitative, descriptive survey design was employed using a validated 21-item questionnaire administered to 350 full-time faculty members at Al-Zaytoonah University of Jordan. The instrument demonstrated strong reliability (Cronbach's $\alpha = 0.91$) and sound construct validity confirmed through confirmatory factor analysis (CFI = 0.95, RMSEA = 0.06). Results indicated a high overall level of acceptance of IoT applications in teaching ($M = 4.12$, $SD = 0.88$). No statistically significant differences were found by gender, while small but statistically significant differences emerged by academic rank, with assistant and associate professors reporting more positive perceptions than full professors ($\eta^2 = 0.024$). The findings suggest that IoT acceptance is broadly shared among faculty, with academic rank functioning as a modest, context-dependent moderator. The study contributes empirical evidence on IoT-enabled e-learning practices in Middle Eastern private higher education and highlights the need for targeted professional development and institutional support strategies.

Keywords: Internet of things (IoT), Faculty perceptions, Academic rank, Technology adoption, Jordan

1. Introduction

Technology has brought great changes to the global education sector, as it has transformed how we teach, assess and learn. One of the key technological achievements that emerges from this process is the Internet of Things, whose physical and digital worlds are "networked to allow numerous devices and sensors to interact with and access a variety of systems and data". In addition, this leads to further deployment of smart classrooms, real-time assessment, enhanced communication, and other spaces, improving the educational landscape of universities (Villegas-Ch, Palacios-Pacheco, & Román-Cañizares, 2020). In part, this is the reason why to compete in the labor market and the future of work (Mershad & Wakim, 2018).

Having faculty members who are able to implement and deploy IoT in all universities will be key. It thus influences the perceived utility of technology and its sustained use for further learning. This extends to Technology Adoption Models such as TAM or UTAUT in which the effect of perceived usefulness, perceived ease of use and social influences on technology acceptance are studied (Venkatesh et al., 2003). These constructions exist at the level of higher education through the "institutional culture, professional norms, and organizational expectations…which affect faculty involvement with teaching technologies" (Davis, 1989). They also represent impacts on the core adoption constructs, individual differences, and other factors Gender and rank might also play a role in faculty perception about how useful IoT will be to teach their respective courses, and in the effort, they put into using IoT for teaching.

A few studies on IoT in education have been conducted in the Middle East and elsewhere. In this regard, for example, El-Dahshan (2019) observed some favorable attitudes among faculty toward the IoT in Egypt and Al-

Cite this article: Abu Jraiban, T. et al. 2026. "From Chalkboards to Smart Classrooms: Faculty Perceptions of IoT Integration in Jordanian Universities", *Electronic Journal of e-Learning*, 24(2), pp 82-93, https://doi.org/10.34190/ejel.24.2.4503

Ma'mari et al. (2019) reported on the willingness of participating faculty members to integrate IoT at Sultan Qaboos University in Oman. In the same way, at an international level, it has been found that academics see IoT as useful for administrative efficiency and instructional support (Villegas-Ch, Palacios-Pacheco, & Román-Cañizares, 2020) Yet little is known about faculty attitudes in Jordan towards the use of IoT – especially that of private universities, which are increasingly influential actors in making higher education more accessible. Private universities in Jordan do not have the public authority of their public counterparts, but instead function within a more market-based environment, pitted against one another for students and with limited resources at their disposal. This and the subsequent section are because at the end of the day, in the context of Jordan, the private HEIs are an immensely important and unique place to study the adoption of technology because of the levels of competition and resources available to these universities.

On top of that, the analyses of socio-demographic data—gender or academic rank—should be taken into account when employing these theoretical assumptions. For instance, under UTAUT, gender is noted to modulate social influence and behavioral intention, so it could be investigated in other cultures. Grades may also be a measure of education and experience with facilitating conditions, like institutional support and training. Several senior faculties have more established teaching and incentive structures that ultimately affect their effort expectancy and performance expectancy. Although previous studies did not indicate any significant gender differences in the use of educational technology, equity and access remain important factors in higher education. Higher education is just as important; some international studies do not find any differences in rank between instructors (Mershad & Wakim, 2018). But new research has suggested that higher-level faculty are more likely to adopt digital technologies due to enhanced ICT training and digital education. Sprenger and Schwaninger (2021), for example, stated that perceived usefulness and ease of use constituted most acceptance of digital learning tools. Such findings show the role of individual and institutional factors in adoption behaviors.

Despite international research expanding rapidly, empirical evidence regarding faculty understandings of IoT adoption in private universities in Jordan is absent. This gap challenges the international models of technology acceptance in higher education in Jordan.

The study aims to two major ends: to first look to how faculty perceive IoT applications in university teaching and how they are useful in learning and in academics. Second, to understand how the perceptions differ according to gender and level of instruction, exploring how they are moderated by the variables, as suggested by the accepted model of technology acceptance, and how personal and professional characteristics affect faculty's use of new technologies.

RQ1: What are the perceptions of faculty members in Jordanian private universities toward the use of IoT applications in university teaching?

RQ2: Are there statistically significant differences in these perceptions attributable to gender and academic rank?

This is a theoretical and practical study which has of great value. It theoretically applies TAM and UTAUT into a context underexplored at the institutional and regional level via gender and academic rank as moderating variables. In practice, this information can be used to inform faculty development efforts, institutional ICT policy and formulate a plan for digital transformation at Jordan's private higher education.

2. Literature Review

In this review I have analyzed the major theoretical models of technology adoption, the perceived utility and simplicity of IoT, gender and rank, and identified the research gap in Jordanian private universities.

2.1 Theoretical Foundations: TAM and UTAUT in Educational Technology

The underlying idea of this study follows two leading models of technology acceptance. First, the TAM, which argues that the intention to use a technology is based mostly on two things: perceived usefulness (the extent to which someone thinks that the application of a system will improve their performance) and perceived ease of use (the extent to which people think that the application of a system will be effort-free) (Davis, 1989). Building on TAM, the Unified Theory of Acceptance and Use of Technology UTAUT considers four direct factors related to intention and use: performance expectancy; effort expectancy; social influence; and facilitating conditions (Venkatesh et al., 2003). UTAUT also notes that the strength of these relationships is influenced by the individual traits of gender and experience, such as in a context of technology acceptance.

These models are especially relevant to higher education, because faculty decision-making emerges within institutional culture, professional practices and organizational expectations, providing an attractive way of investigating IoT integration in university instruction. in university instruction.

2.2 Perceived Usefulness and Applications of IoT in Education

The perceived benefits of the Internet of Things in higher education have been explored in many ways. As with many studies, IoT provides access to smart classrooms, interactive learning environments, and real-time feedback systems to provide individual and flexible instruction (Jayousi et al. 2025; Alaklabi, 2019). Not only in teaching, but also IoT has been linked to better communication, coordination, and administration at universities (Shihao, Dahnil, & Saad, 2025; Suster et al. 2025).

Matar Al-Salmi, Abdullah, and Al-Hinai (2020) explored the impact of IoT in allowing for management of learning and how this technology could change the way information is stored, used in schools. Like IoT applications, IoT applications have been associated with improved academic support services, such as access to library resources and digital infrastructures (Saha & Roknuzzaman, 2024). International reports support these results Villegas-Ch, Palacios-Pacheco, and Román-Cañizares (2020) found that Ecuadorian faculty perceived the value of IoT to improve administrative efficiency and Mershad and Wakim (2018) found that IoT enhanced students' flexibility and teaching time.

Recently, studies suggest that experiential and practical experiences with IoT technologies can also have an effect on acceptance among teachers and students, particularly as IoT is embedded in instruction (Varela-Aldás et al. 2025). In addition, Abdelhamid (2021) found that smart learning environments built on IoT supported smart learning systems were especially helpful in improving digital literacy and student-teacher understanding of technology. A second key component of the adoption of IoT is perceived to be utility, but embodies classroom integration, not technological availability.

2.3 Perceived Ease of Use and Barriers to IoT Adoption

While IoT benefits are great, costs and effort remain barriers to adoption and the effects of international and country-specific research experience often diminish perceived ease of use, practical challenges are also preventing people from being able to easily access the Internet. Barriers include privacy and data security issues, limited infrastructure, high implementation costs and poor institutional training (El-Dahshan, 2019; Al-Ma'mari et al., 2019). Further, IoT requires that institutions handle large volumes of data, making for further complications of knowledge management systems and organizational readiness (Matar Al-Salmi, Abdullah, & Al-Hinai, 2020).

Study after study finds that perceived utility is highly correlated with persistent implementation barriers. For example, in Oman, strong faculty motivation and infrastructure issues are found (Al-Ma'mari et al., 2019); in Saudi Arabia, usability, institutional support, and security concerns are found to be top reasons for adoption (Ali, Syed, & Danish, 2023). Also, research from developing countries points out that institutional readiness and facilitators are the most important factors critical for IoT adoption (Madni et al., 2022). These findings indicate that ease of use is not just a technical problem but a systemic problem, driven by organizational support and policy.

2.4 The Role of External Variables: Gender and Academic Rank in Technology Adoption

Although the main constructs of TAM and UTAUT are universally applicable, they also provide insights into. Gender and class are contextual factors of technology adoption that should be examined more closely in higher education.

Gender. In the UTAUT model gender is involved in modulating the link between social power, effort expectancy, and behavioral intent. The empirical results are mixed. Although some studies do not reveal significant gender differences in the perception of technology when access and support are equitable (Villegas-Ch, Palacios-Pacheco, & Román-Cañizares, 2020). Nonetheless, given the persistent challenges to equity and digital inclusion within Middle Eastern higher education, exploring gender has less philosophical meaning than it should.

Academic rank. A higher level of education may provide a measure of age, experience, and learning of new instructional practices. However, new research suggests that recent ICT training and higher support for teaching innovation can lead junior faculty toward digital adoption (Ali, Syed, & Danish, 2023). Conversely, senior faculty might depend on established teaching practices and have their own professional incentive structures. Nevertheless, the findings are undoubtedly inconsistent; Mershad and Wakim (2018) reported no rank difference between U.S. institutions. These contradictions suggest that the role of academic rank has an integral contextual dimension and is mediated by institutional culture, policy making, and available resources. Academic Status. The higher education level may provide a measure of age, learning, and experiences with new instructional practices. But new research suggests that recent ICT training and higher support for teaching innovation can lead junior faculty toward digital adoption (Ali, Syed, & Danish, 2023). The senior faculty, on the other hand, may use established teaching practices and have their own professional incentives. But the results were clearly inconsistent; Mershad and Wakim (2018) reported no difference in rank between U.S. institutions. These contrasts show that the influence of professorship is highly contextual and dependent upon institutional culture, governance system, and available resources.

2.5 Synthesis and Identified Research Gap

The literature supports TAM and UTAUT as an effective model for the study of IoT adoption with perceived utility and ease of use being cited as main factors. Few studies have examined the relationship between demographic and institutional context in defining adoption, particularly in private Jordanian universities. These institutions have competitive market pressures and resource constraints and a unique circumstances under which technology is used.

This clarifies what should be done in terms of empirical research regarding the perception of faculty regarding the adoption of IoT across gender and class in private education in Jordan. This gap provides context for discussion and for regional scholarship and further discourse regarding the limits of universal models of technology acceptance.

3. Methodology

This study employed a quantitative, descriptive survey design to investigate faculty perceptions of IoT applications in teaching at Al-Zaytoonah University of Jordan, a representative private university in Jordan. The study was conducted during the second semester of the 2023/2024 academic year and was limited to full-time faculty members. Although the single-institution design limits generalizability, it was selected to allow for a deep, controlled examination of the institutional culture and demographic variables within a defined setting, providing a robust case study for the Jordanian private higher education sector. The scope of the study is therefore bounded by this institutional and temporal context. Other limitations, including the use of self-reported data and the cross-sectional design, are discussed in a later section.

3.1 Research Design

The descriptive, quantitative survey design was used because it is well suited to analyzing attitudes and perceptions over a particular population at a given point in time (Creswell & Creswell, 2018). This design provides a means of rigorously examining faculty perceptions and demographic group comparisons, as well as consistent with the methodology commonly used in technology acceptance studies (Madni et al., 2022; Ali, Syed, & Danish, 2023). There were no directional hypotheses, but the study follows an analytical logic based on theory based on TAM and UTAUT that looks at gender and rank as moderating variables that may change faculty perceptions of IoT adoption.

Figure 1 presents the conceptual model of the study, illustrating the theoretical framework grounded in TAM and UTAUT and the role of gender and academic rank as contextual moderating variables influencing faculty perceptions of IoT adoption in teaching.

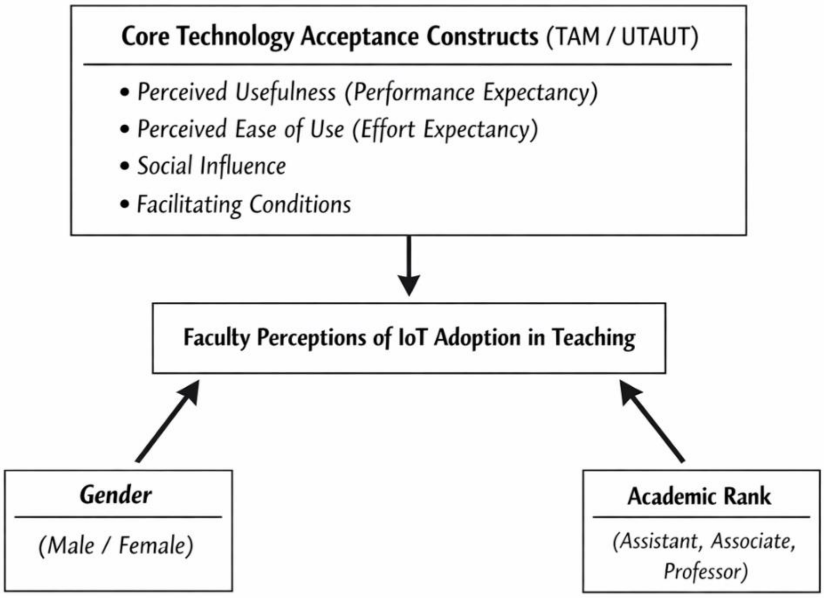


Figure 1: Conceptual model of faculty perceptions of IoT adoption based on TAM and UTAUT

3.2 Population and Sample

All 741 full-time faculty members were included in the study population at Al-Zaytoonah University of Jordan.

Generally, a minimum representative sample size of 253 was recommended in Krejcie and Morgan (1970) table.

To provide sufficient statistical strength and to minimize of non-response bias, a higher number of questionnaires were distributed.

- Distribution: A total of 357 questionnaires were distributed: 304 in paper format and 53 online via Google Forms.
- Collection and Exclusion: All 304 paper questionnaires were returned. Four were excluded due to extensive missing data, yielding 300 valid paper responses. All 53 online responses were collected; three were excluded for the same reason, yielding 50 valid online responses.
- Final Sample: The final valid sample for analysis was therefore N = 350 faculty members. The demographic distribution is presented in Table 1.

Table 1: Distribution of the Study Sample by Gender and Academic Rank (N = 350)

Variable	Category	n	%
Gender	Male	211	60%
	Female	139	40%
Academic rank	Professor	114	33%
	Associate Professor	126	36%
	Assistant Professor	110	31%

3.3 Instrument Development and Description

A structured questionnaire was developed to assess faculty perceptions based on the core constructs of the Technology Acceptance Model (TAM) and the Unified Theory of Acceptance and Use of Technology (UTAUT). The initial 25-item pool was generated through a review of the literature on IoT adoption in higher education (e.g., Abdelhamid, 2021; Madni et al., 2022). The final instrument operationalized four theoretical constructs:

Perceived Usefulness / Performance Expectancy (6 items)

Perceived Ease of Use / Effort Expectancy (5 items)

Social Influence (4 items)

Facilitating Conditions (6 items)

Content and face validity were established through expert review by eight specialists in educational technology and measurement. Based on their feedback, four overlapping items were removed, resulting in a final instrument of 21 items. Responses were measured using a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree).

3.4 Validity and Reliability

They conducted a small-scale research project with 60 faculty members in order to evaluate their psychometric properties. Construct Validity: Confirmatory Factor Analysis (CFA) showed the proposed four-factor structure. The model fit indices were excellent: $\chi^2/df = 2.15$, CFI = 0.95, TLI = 0.93, RMSEA = 0.06, and SRMR = 0.04, all meeting thresholds for good fit (Hu & Bentler, 1999).

Standardized factor loadings for all items were in the range of 0.63 to 0.87, exceeding the minimum recommended of 0.60, suggesting strong convergent validity at the item level. No item had cross-loadings or weak structural relationships to the intended construction and all loadings were statistically significant ($p > .001$).

For each construct, convergent validity was further tested with (AVE) and (CR). The AVE values were 0.54–0.86 exceeding the desired cutoff of 0.50, and the CR values were 0.83–0.91 higher than the suggested cutoff of 0.70. This reveals a significant component of the variance in indicators within each construct.

The discriminant validity was assessed with the Fornell–Larcker criterion. This is explained by an observation that AVE is larger than the inter-construct correlation coefficients for all constructs so that each construct is distinct and reflects a different dimension of faculty perception of IoT adoption.

The Cronbach's alpha of the instrument was 0.91, showing strong internal consistency. The subscales also showed strong reliability: Perceived Usefulness ($\alpha = .88$), Perceived Ease of Use ($\alpha = .85$), Social Influence ($\alpha = .82$), and Facilitating Conditions ($\alpha = .87$). The correlation coefficient for the test-retest reliability was 0.89 within two weeks, which indicated a strong correlation ($p < .001$).

3.5 Operationalization of Variables

While gender, female and class, were independent variables, male and female variables were both employed. assistant, associate and full professor variables. Academic rank was also viewed as an indicator of professional prestige and affiliation, instead of age or experience. In this case age or training did not matter, as explained in the discussion section.

3.6 Ethical Considerations

Participation was voluntary and an overview was given to each participant before data collection to clearly articulate the purpose of the study. Informed Consent was obtained with strict confidentiality and anonymity. Reporting was done in aggregate.

3.7 Procedures and Bias Mitigation

Data collection was conducted according to a systematic format. Independent-sample t-tests between early and late respondents were used to test for bias by non-response bias and no significant differences were observed ($p>.05$). Procedural remedies were used to limit common method bias, such as assurance of anonymity and reverse-coded items. Statistically, Harman's one factor test found that no single factor contributed to the majority of variance (38.2%), but well below the 50% recommended limit.

3.8 Statistical Analysis

All data was used for analysis with SPSS v27 and AMOS v24. Research Question 1 was assessed using descriptive statistics (means, standard deviations, and confidence intervals). We conducted independent-samples t-tests for gender differences and one-way ANOVA using Tukey's HSD post hoc tests to test for gender differences. The effects sizes, Cohen's d and eta squared, were reported in systematic ways to supplement the statistical significance testing and make it more interpretable, as required in educational research.

4. Results and Findings

The findings are presented as three sections: perceptions of IoT applications among faculty; gendered and rank-based variations; and summary of the findings.

4.1 Research Question 1: Faculty Perceptions of IoT Applications

Descriptive statistics were then used to measure the general perception of faculty members regarding IoT applications in higher education teaching. Table 2 shows that faculty members reported a high level of positive perception overall, with a mean score of M = 4.12 (SD =0.88) on a 5-point Likert scale. Additionally, the 95% confidence interval of [4.04, 4.20] indicates that the population mean is in the upper agreement range, which implies that IoT is a strongly supported teaching practice.

It found the highest mean scores for statements on IoT's role in real-time student assessment (Item 21, M = 4.47), enhanced instruction through multimedia (Item 19, M = 4.46), and traditional teaching methodologies (Item 20, M = 4.44). These items are reflective of IoT's perceived instructional value and pedagogical utility, not technology's singular appeal.

Table 2: Faculty Members' Perceptions of IoT Applications in University Teaching (N = 350)

Rank	Item No.	Item Statement	Mean	SD	95% CI	Perception Degree
1	21	IoT enables teachers to measure students' learning progress in real time.	4.47	0.91	[4.37, 4.57]	Very High
2	19	IoT can be used to teach a wide range of subjects using graphics and animations to enhance understanding.	4.46	1.02	[4.34, 4.58]	Very High
3	20	IoT allows for a transformation from traditional to digital teaching methodologies, with increased efficiency.	4.44	0.79	[4.34, 4.54]	Very High
4	18	IoT devices provide reliable access to educational materials and communication channels.	4.41	0.89	[4.30, 4.52]	High
5	15	IoT contributes to scientific research and the formation of research collaboration groups.	4.40	0.75	[4.32, 4.48]	High
6	16	IoT provides diverse teaching methods.	4.38	0.88	[4.28, 4.48]	High
7	17	IoT helps simplify the explanation of concepts and link experiences to simulated reality.	4.36	0.89	[4.26, 4.46]	High
8	10	IoT saves instructors' time through faster attendance registration and assignment collection.	4.32	0.81	[4.23, 4.41]	High
9	14	IoT strengthens student–teacher connections through direct communication in virtual classrooms.	4.30	0.99	[4.19, 4.41]	High
10	13	IoT assists in students' comprehension by diversifying instructional strategies.	4.28	0.76	[4.20, 4.36]	High
11	12	Adequate infrastructure for IoT facilitates the teaching–learning process.	4.25	0.92	[4.15, 4.35]	High
12	11	The use of IoT applications is easy and straightforward.	4.22	0.85	[4.13, 4.31]	High
13	1	IoT reduces burdens on students.	4.21	1.00	[4.10, 4.32]	High
14	5	IoT use does not pose fears of security breaches, surveillance, or similar threats.	4.20	0.92	[4.10, 4.30]	High
15	4	IoT use leads to comprehensive experience with the latest technologies.	4.12	0.78	[4.04, 4.20]	High
16	3	IoT provides empowering tools to access organizational and technological knowledge.	4.10	0.82	[4.02, 4.18]	High
17	2	IoT use develops new skills and competencies in information technology.	4.01	0.91	[3.91, 4.11]	High
18	6	IoT supports innovation and reduces process execution time.	3.89	0.93	[3.79, 3.99]	Moderate
19	9	IoT provides a rich and flexible platform to explore learning in an intelligent environment.	3.88	0.85	[3.79, 3.97]	Moderate
20	8	IoT helps eliminate repetitive daily tasks and focus on more important matters.	3.84	1.00	[3.73, 3.95]	Moderate

Rank	Item No.	Item Statement	Mean	SD	95% CI	Perception Degree
21	7	IoT supports learning anytime and anywhere.	3.74	0.92	[3.63, 3.85]	Moderate
		Overall Perception	4.12	0.88	[4.04, 4.20]	High

Note. CI = Confidence Interval. Scale range = 1 (strongly disagree) to 5 (strongly agree).

4.2 Research Question 2: Differences by Gender and Academic Rank

4.2.1 Gender differences

In order to compare the perceptions of female and male faculty members, a t-test using independent-samples was undertaken. As reported in Table 3, there was no statistically significant difference in faculty perception between males and females, $t(348) = 1.18$, $p = .24$. Cohen's d calculated the effect size as 0.13 and therefore extremely small. This result indicates that gender is not at all involved in the difference between faculty's perceptions of IoT adoption.

Table 3: Independent-Samples t-Test for Gender Differences in Perceptions

Gender	n	M	SD	t(348)	p	Cohen's d	95% CI (diff)
Male	211	53.28	7.00	1.18	.24	0.13	[–0.09, 0.35]
Female	139	52.11	6.84				

4.2.2 Differences by academic rank

To examine variations in how faculty groups perceive their institutions, a one-way analysis of variance was used to compare faculty views across academic rank. A statistically significant differentiation between the ranks was found, $F(2,347) = 4.27$, $p = .015$, $\eta^2 = .024$ (see Table 4). The large sample size may contribute to that this value is statistically significant, although this effect size is so small that it is likely meaningless in practical terms.

Table 4: One-Way ANOVA for Perceptions by Academic Rank

Source	SS	df	MS	F	p	η²
Between groups	321.25	2	160.63	4.27	.015	0.024
Within groups	13,020.6	347	37.52			
Total	13,341.9	349				

This small effect size suggests that academic rank is a statistically significant but practically insignificant factor in explaining differences in faculty attitudes toward IoT adoption.

To determine the source of these differences, after the fact comparisons were performed using Tukey's HSD test (Table 5).

Table 5: Tukey Post Hoc Comparisons of Academic Rank

Comparison	Mean Diff.	p	95% CI
Professor – Associate Prof.	–0.26	.09	[–0.56, 0.04]
Professor – Assistant Prof.	–0.30*	.03	[–0.58, –0.02]
Associate – Assistant Prof.	–0.56*	.00	[–0.85, –0.27]

Post hoc estimates indicated that assistant professors scored far higher on positive perception than full professors ($p = .03$) and associate professors ($p < .001$) and were not statistically significant in comparison to professors and associate professors ($p = .09$). The differences in mean were small, though statistically significant, and support the notion that academic rank serves as a secondary rather than a major influence on perceptions.

In other words, the mean is 0.30 to 0.56 on the 5-point scale, indicating relatively little similarity in practical instruction across faculty.

4.3 Summary of Findings

In general, the faculty indicated that they were embracing IoT applications in their teaching practice. Gender-related differences were insignificant, although significant differences were observed across academic rank. Importantly, these rank-based differences were minor effects, suggesting that academic rank is only one of many factors that interact with faculty perceptions rather than being the main driver.

5. Discussion

These findings directly address the first research question on faculty's perceptions of IoT in teaching. Faculty at private Jordanian universities shared broadly positive sentiment, with scores particularly high on IoT's effectiveness in making real-time assessments of students, regarding resources and in changing teaching practices. These findings are consistent with other studies indicating that IoT enhances school efficiency, instructional flexibility and higher learning outcomes (Abdelhamid, 2021; El-Dahshan, 2019; Mershad & Wakim, 2018). Importantly, the consistently high mean scores across items suggest not a specific acceptance of a particular tool, but a wider perception of IoT as a teaching paradigm rather than as a supplementary technology. The benefits reported are strongly associated with the core constructs of Technology Acceptance Model (TAM) and the Unified Theory of Acceptance and Use of Technology (UTAUT). High levels of perceived usefulness (performance expectancy) and manageable effort expectancy indicate that faculty interpret IoT as contributing to teaching effectiveness without too much cognitive or technical pressure (Venkatesh et al., 2003). This balance of perceived effort with anticipated performance improvement seems essential for acceptance of private universities with IoT in financial constraints, and with instructional efficiency at the heart of resource-constrained private colleges in the cloud. These results converge with Sprenger and Schwaninger (2021), who found that perceived usefulness and ease of use influence adoption goals across digital learning technologies. Also found was an answer to the second question the study, demographic differences. This finding corroborates prior research that based on a similar combination of access, infrastructure, and institutional support, technology is increasingly gender-neutral in higher education (Villegas-Ch, Palacios-Pacheco, & Román-Cañizares, 2020). This result suggests that social influence and the supportive conditions of employment at UTAUT are similar to those of male and female faculty in Jordanian private universities, reflecting progress toward more equitable institutional engagement with digital resources over gendered adoption.

By contrast, the statistically significant, but almost invisible difference in academic rank was the difference between assistant and associate professors, both of whom expressed greater positive views than full professors. The small effect size ($\eta^2 = 0.024$) indicates that school rank accounts for a tiny proportion of variance in IoT perceptions, and thus rank is not a dominant factor in adoption attitudes. But, if the difference is statistically significant, it should be contextualized rather than dismissed.

Junior teachers may have more favorable perceptions in terms of the performance expectancy of assessment, student learning, and job recruitment requirements. Rather than being solely generational, these findings are more readily understood through professional incentives and institutional role expectations. Assistant professors, particularly, are more inclined to use new instructional methods, and associate professors may blend teaching with promotion-related needs. Full professors, on the other hand, can operate on more stable professional routes that diminish the perceived marginal gain associated with using new instructional technologies, even with institutional support.

These patterns may also reflect differences in the facilitation conditions. Higher teachers' workload for junior faculty may contribute to increased perceived value for IoT tools that improve assessment, communication, and instructional management, thus improving effort expectancy by integrating performance with effort. On top of that, professors who have received doctoral training in rapid ICT integration are more likely to have taught digital classes and to be less stressed when IoT adoption is evident.

The presence of rank-based differences is not consistent with those found in other institutional settings (Al-Ma'mari et al., 2019; Mershad & Wakim, 2018), which suggests that institutional context is the important moderator for technology acceptance. In private Jordanian institutions, pressures created by market forces, performance-based assessment systems, and student satisfaction may further strengthen the perceived instructional value of IoT for faculty members actively developing or integrating their profiles. This contextual sensitivity further demonstrates the limitations of treating these models of technology acceptance as uniformly applicable across institutional domains.

Interpretation of these results should be accompanied by several limitations. The study was conducted at one private university, which limited generalizability. This approach may result in response bias from the self-reported perceptions and longitudinal aspects in adoption attitudes cannot be examined because the design is cross-sectional. Furthermore, academic rank was used as a measure of professional experience and career stage without explicit measures of age, teaching load or research expectations. Future studies should adopt mixed-methods, incorporate behavioral characteristics of IoT use and look at multiple institutional structures in order to better understand the mechanisms underlying demographic differences in technology acceptance.

6. Conclusion

The purpose of this study was to understand the perceptions of university personnel toward the implementation of the IoT at private Jordanian universities and how they differ across gender and socioeconomic status. The results suggest that the impact of IoT is strongly acknowledged by faculty as facilitating effectiveness of teaching, student participation, and instructional support. This lack of gender-based differences reflects a similar level of access to and willingness to use digital technologies among faculty. While statistically significant differences by academic rank were found, their practical magnitude was negligible, suggesting that rank is more a contextual rather than a primary determinant of IoT acceptance. More specifically, associate and assistant professors had more positive perceptions than full professors; however, the small effect size suggests that academic rank is one of several relationships that influence attitudes about technology.

The primary outcome of this study is to show the influence of institutional context on technology acceptance. In the private universities of Jordan with market-driven governance and variable allocations of resources, academic rank seems to be intertwined with professional expectations and incentives, thus producing minor differences in faculty perceptions. This contradicts data in other national and institutional settings, and illustrates the context-sensitive nature of models of technology acceptance. Therefore, their research presents TAM and UTAUT as reflecting the nature of core constructs in local institutional bodies rather than a universally invariant system.

This study has academic and policy implications. If IoT integration is effective, it must be coordinated across institutions, not one faculty effort at a time. In the university, for instance, structured peer-mentoring may be possible, for instance, for sharing knowledge at all stages of a career. At policy level, the systemic involvement of digital pedagogy and IoT readiness in quality assurance and accreditation could encourage further adoption. Furthermore, faculty development programs should be shifted to support senior staff through training, innovation grants or teacher load adjustment so that the small difference in rank considered in this study is not considered in the programs.

There are a few limitations. This is not the single private university that is difficult to generalize, and there is no use for cross-sectional data for causal inference. Future research studies in this area should be conducted in mixed methods and longitudinal studies designed to evaluate the impact professional experience, educational burden, and institutional incentives can have on technology adoption over time. Comparisons between public and private universities and national government would further show how governance and resource structures are linked.

Finally, this paper reflects the prevalence of high levels of acceptance of IoT at private Jordanian universities and has some observations on small contextual variation regarding the relationship between ranking and status. In addition, in considering demographics and institutional context, this study provides new insight into a more complex literature on technology acceptance and offers useful insights for education policy and practice for the development of higher education systems.

To conclude, this study provides empirical evidence for significant faculty acceptance of IoT in Jordanian private universities, but presents minor, contextually dependent variation relative to rank. The study, which integrates demographic context with institutional context, provides a new perspective on the literature in acceptance of technology and provides useful insights into educational policies and practice regarding improving the higher education system.

Conflict of Interest: The authors acknowledge that any commercial, financial or personal relationships that could be seen as possible conflicts of interest with respect to performance or publication of this study are irrelevant.

AI Statement: The authors declare that no artificial intelligence tools were used at any stage of the research process or in the preparation, writing, editing, or revision of this manuscript. All content was developed entirely by the authors.

Ethics Statement: Ethical approval for this study was obtained from the Ethics Committee of Al-Zaytoonah University of Jordan (Approval No. MAR/24). All participants were informed of the purpose of the research prior to

data collection and participated voluntarily. Informed consent was obtained from all respondents. Participation was anonymous, and no personally identifiable information was collected. All data were treated with strict confidentiality and reported in aggregate form only. The study was conducted in accordance with the ethical principles of the Declaration of Helsinki.

References

Abdelhamid, R.H. (2021) 'Effectiveness of a proposed program in light of the requirements of the Fourth Industrial Revolution using a smart learning environment based on the Internet of Things to develop digital teaching skills, future foresight, and technological acceptance among female mathematics student-teachers', *Journal of Mathematics Education*, 24(1), pp. 182–267. Available at: https://doi.org/10.21608/armin.2021.143152 (Accessed: 14 October 2025).

Alaklabi, A.T. (2019) 'The feasibility of Internet of Things applications in education', *International Journal of Research in Educational Sciences*, 2(3), pp. 147–166. Available at: https://iafh.net/index.php/IJRES/article/view/122 (Accessed: 14 October 2025).

Ali, J., Syed, M. and Danish, A. (2023) 'IoT adoption model for e-learning in higher education institutes: a case study in Saudi Arabia', *Sustainability*, 15(12), 9748. Available at: https://doi.org/10.3390/su15129748 (Accessed: 14 October 2025).

Al-Ma'mari, A., Al-Kindi, A., Al-Dhahli, M. and Al-Farsi, H. (2019) 'Technological acceptance of the Internet of Things in the educational process in the Department of Information Studies at Sultan Qaboos University'. In: *Proceedings of the 25th Annual Conference of the Special Libraries Association – Arabian Gulf Chapter: Internet of Things: The Future of Interconnected Internet Societies*. Abu Dhabi, 5–7 March. Available at: https://search.mandumah.com/Record/946680 (Accessed: 14 October 2025).

Creswell, J.W. and Creswell, J.D. (2018) *Research design: qualitative, quantitative, and mixed methods approaches*. 5th edn. Thousand Oaks, CA: SAGE Publications.

Davis, F.D. (1989) 'Perceived usefulness, perceived ease of use, and user acceptance of information technology', *MIS Quarterly*, 13(3), pp. 319–340. Available at: https://doi.org/10.2307/249008 (Accessed: 14 October 2025).

El-Dahshan, G. (2019) 'Employing Internet of Things in education', *International Journal of Research in Educational Sciences*, 2(3), pp. 49–92. Available at: https://iafh.net/index.php/IJRES/article/view/121 (Accessed: 14 October 2025).

Hu, L. and Bentler, P.M. (1999) 'Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives', *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), pp. 1–55. Available at: https://doi.org/10.1080/10705519909540118 (Accessed: 14 October 2025).

Jayousi, S., Lucattini, P., Petti, L., Bruni, F. and Mucchi, L. (2025) 'Smart learning by design: a framework for IoT-driven adaptive classrooms and inclusive education', *Education Sciences*, 15(10), 1338. Available at: https://doi.org/10.3390/educsci15101338 (Accessed: 14 October 2025).

Krejcie, R.V. and Morgan, D.W. (1970) 'Determining sample size for research activities', *Educational and Psychological Measurement*, 30(3), pp. 607–610. Available at: https://doi.org/10.1177/001316447003000308 (Accessed: 14 October 2025).

Madni, S.H.H., Ali, J., Husnain, H.A., Masum, M.H., Mustafa, S., Shuja, J., Maray, M. and Hosseini, S. (2022) 'Factors influencing the adoption of IoT for e-learning in higher educational institutes in developing countries', *Frontiers in Psychology*, 13, 915596. Available at: https://doi.org/10.3389/fpsyg.2022.915596 (Accessed: 14 October 2025).

Matar Al-Salmi, J.B., Abdullah, K.A.S. and Al-Hinai, A.B.S. (2020) 'The role of the Internet of Things in knowledge management in information institutions', *Journal of Information Studies and Technology*, 2020(1), 3. Available at: https://doi.org/10.5339/jist.2020.3 (Accessed: 14 October 2025).

Mershad, K. and Wakim, P. (2018) 'A learning management system enhanced with Internet of Things applications', *Journal of Education and Learning*, 7(3), pp. 23–40. Available at: https://doi.org/10.5539/jel.v7n3p23 (Accessed: 14 October 2025).

Saha, S. and Roknuzzaman, M. (2024) 'Library practitioners' perceptions on the applications of IoT in university libraries of Bangladesh', *Library Management*, 45(1/2), pp. 141–156. Available at: https://doi.org/10.1108/LM-07-2023-0072 (Accessed: 14 October 2025).

Shihao, L., Dahnil, D.P. and Saad, S. (2025) 'A survey of smart campus resource information management in Internet of Things', *IEEE Access*, 13, pp. 66622–66645. Available at: https://doi.org/10.1109/ACCESS.2025.3558900 (Accessed: 14 October 2025).

Sprenger, D.A. and Schwaninger, A. (2021) 'Technology acceptance of four digital learning technologies after three months' usage', *International Journal of Educational Technology in Higher Education*, 18(1). Available at: https://doi.org/10.1186/s41239-021-00243-4 (Accessed: 14 October 2025).

Suster, G., Popescu, C.A., Iancu, T., Popescu, G. and Ciolac, R. (2025) 'The synergy of smart campus development with smart city policies and the New European Bauhaus with implications for educational efficiency', *Sustainability*, 17(17), 8078. Available at: https://doi.org/10.3390/su17178078 (Accessed: 14 October 2025).

Varela-Aldás, J., Junta, C., Choque, E. and Palacios-Navarro, G. (2025) 'Influence of higher education on IoT acceptance through hands-on learning', *TEM Journal*, 14(1), pp. 528–539. Available at: https://doi.org/10.18421/TEM141-47 (Accessed: 14 October 2025).

Venkatesh, V., Morris, M.G., Davis, G.B. and Davis, F.D. (2003) 'User acceptance of information technology: toward a unified view', *MIS Quarterly*, 27(3), pp. 425–478. Available at: https://doi.org/10.2307/30036540 (Accessed: 14 October 2025).

Villegas-Ch, W., Palacios-Pacheco, X. and Román-Cañizares, M. (2020) 'An Internet of Things model for improving process management on university campus', *Future Internet*, 12(10), 162. Available at: https://doi.org/10.3390/fi12100162 (Accessed: 14 October 2025).

Digital Pedagogy in Indian Higher Education: Faculty Perspectives

Divneet Kaur[1], Shikha Rana[2], Nishant Chaturvedi[3], Shalini Bahuguna Bachheti[4] and Vinay Punia[5]

[1]School of Business, UPES, Dehradun, India
[2]School of Management, Graphic Era Hill University, India
[3]Swami Rama Himalayan University Dehradun, India
[4]ICFAI Law School, ICFAI University, Dehradun, India
[5]Amrapali University- Haldwani, India

divneetbagga04@gmail.com
shikharana.ddn@gmail.com (corresponding author)
fdp17nishantc@alumni.iimidr.ac.in
shalinibahuguna@iudehradun.edu.in
vnypunia@gmail.com

https://doi.org/10.34190/ejel.24.2.4189

Abstract: The rapid digitalization of higher education has significantly reshaped teaching and learning practices worldwide; however, the adoption of digital pedagogy among university teachers remains uneven, particularly in developing contexts such as India. This study examines the lived experiences of Indian university teachers in adopting digital pedagogy and explores the factors influencing this process within higher education institutions. Using a qualitative research design, the study employs Interpretative Phenomenological Analysis to develop an in depth understanding of how teachers perceive, experience, and make sense of digitally mediated teaching practices. Data were collected through semi structured interviews with university teachers representing diverse disciplinary backgrounds and institutional settings. The analysis followed a systematic and iterative IPA approach to identify emergent themes grounded in participants' narratives. The findings indicate that digital pedagogy adoption is shaped by a dynamic interplay of institutional, technological, and personal factors. Institutional support structures, availability of digital infrastructure, access to professional development opportunities, and collaborative peer environments emerged as key enablers of adoption. In contrast, challenges such as inadequate training, inconsistent technical support, increased workload, infrastructural disparities between institutions, and varying levels of digital confidence among teachers were identified as persistent barriers. The study further highlights the central role of teachers' beliefs, attitudes, and perceived pedagogical value of digital tools in determining the depth and sustainability of digital pedagogy integration. By foregrounding faculty perspectives, this research contributes to the limited qualitative literature on digital pedagogy adoption in Indian higher education and extends existing scholarship beyond technology acceptance oriented explanations. The study supports e learning practice by offering context specific recommendations related to faculty training, institutional policy, and digital readiness. By emphasizing teachers' lived experiences, the findings advance understanding of digital pedagogy as a socially situated and contextually embedded practice, providing a foundation for inclusive and sustainable digital transformation in higher education. The insights generated offer important implications for higher education leaders and policymakers seeking to strengthen digital capacity and enhance teaching quality in evolving educational environments.

Keywords: Digital pedagogy, Interpretative phenomenological analysis (IPA), Higher education sector, Quality assurance, University teachers

1. Introduction

Digital technology in the classroom refers to the use of tools, software, and devices that support learning for all students, including those with specific accessibility needs. It has recently become essential for student development and national growth. In line with the "Digital India" initiative, the government aims to turn the country into a digitally empowered, knowledge-driven society. Reflecting this vision, India's higher education sector, which is the third largest in the world, with over 51,000 institutions and nearly 40 million students, is actively adopting digital transformation. India's Gross Enrolment Ratio (GER) in higher education is 28.4% (AISHE, 2021–22), which is relatively low when viewed against the National Education Policy (NEP) 2020 target of achieving a GER of 50% by 2035 (Government of India, 2020). This level is also below the global average GER of approximately 38% (UNESCO Institute for Statistics, 2022) and significantly lower than that of many developed economies, where GER typically exceeds 60–80% (OECD, 2021; World Bank, 2022 a). Further, access to digital infrastructure is uneven. About 24% of rural households have internet access, while 67% of urban households do (National Family Health Survey-5, 2019-21), highlighting a stark rural–urban disparity in comparison to

Cite this article: Kaur, D. et al. 2026. "Digital Pedagogy in Indian Higher Education: Faculty Perspectives", *Electronic Journal of e-Learning*, 24(2), pp 94-111, https://doi.org/10.34190/ejel.24.2.4189

international benchmarks (World Bank, 2022 b). This disparity creates challenges for adopting digital teaching fairly across different regions. However, the National Education Policy (NEP) 2020 strongly supports integrating digital teaching methods in schools and universities. While digital tools had started to gain popularity before, their use became widespread during the COVID-19 pandemic, accelerating a digital revolution. Technology has since changed the education landscape, affecting students, educators, administrators, and policymakers.

Digital Pedagogy (DP) involves the utilization of contemporary digital technologies in teaching and learning exchange, and it involves reading, accessing, retrieving, and reacting to course materials on digital platforms and devices (Croxal & Kho, 2012; Morris, 2014). Due to the challenges posed by the COVID-19 pandemic and the initiatives of NEP 2020, digital pedagogy has become imperative in the educational environment. The promotion and integration of digital pedagogy in Indian Universities is greatly influenced by the statutory and regulatory bodies, such as the University Grants Commission (UGC) and the All India Council for Technical Education (AICTE), with an aim of augmenting the experience of teachers and students. Scholars have listed various advantages of digital pedagogy, for instance, facilitating the personalized learning experience, enhancing student engagement, and preparing students for digital workplaces (Bećirović, 2023; Schoors et al., 2023). Fuelled by the expectations of students, global competition, and access to a large number of students, it has become significant for teachers to transform their teaching pedagogies (Tom et al., 2023; Montebello, 2017; Sailin & Mahmor, 2018).

The academic community, especially university lecturers, plays an important role in using digital teaching methods. While they have freedom in choosing how to teach, challenges such as limited digital training, heavy workloads, and varying levels of technology readiness make it hard to integrate these tools well. This situation affects student engagement and learning aspects (Bond et al., 2022; Rapanta et al., 2020). As universities prioritize digital transformation to meet evolving student expectations and global educational trends, lecturers' willingness, preparedness, and institutional support become critical for success (Mishra et al., 2020; Tondeur et al., 2017). However, disparities exist in adoption due to infrastructure limitations, discipline-specific needs, and individual digital competence (Bali & Liu, 2018). According to past research, a major challenge for 21st-century teachers is effectively integrating technology into their teaching, which requires strong Technological Pedagogical Content Knowledge (TPACK) or digital pedagogy skills (Ertmer et al., 2012; Milton & Vozzo, 2013). Previously, scholars have opined that technology integration is more likely to be successful if the teacher possesses a constructivist, student-centred pedagogical orientation (Montebello, 2017; Wadmany and Kliachko, 2014). Though the growth of digitalization has propelled the need for using digital pedagogy, the adoption of digital pedagogy is a difficult terrain, especially in an emerging economy like India. Researchers across the globe have identified the various roadblocks in the adoption of digital pedagogy by the academic fraternity for instance; short time span, resources, and technical support (Nanjundaswamy, Baskaran & Leela, 2021), the attitude, beliefs and confidence of teachers in using the digital technologies, their upskilling to match with rapid transformation in DP, perceived support from institutions (Ertmer et al. 2012). These challenges show the need to use e-learning and shift to a more digital approach (Xu et al., 2022). Since digital teaching is still new in India (Efremova & Huseynova, 2022), and considering the differences in internet access and digital infrastructure across Indian regions (National Family Health Survey-5, 2019-21), it is important to examine how university teachers perceive and manage the adoption of digital pedagogy. To gain a deeper understanding of their experiences, this study uses a qualitative method, Interpretative Phenomenological Analysis (IPA). Based on this, the following research questions are proposed:

RQ1 How do Indian university teachers make sense of their lived experiences with adopting digital pedagogy in higher education?

RQ2 What are the key factors influencing university teachers' adoption of digital pedagogy in the Indian higher education sector?

RQ3 What strategies facilitate the effective adoption of digital pedagogy in the HEIs (Higher Educational Institutions), particularly from the perspective of different stakeholders of Indian HEIs?

Considering the above research questions, the study aims to explore the factors affecting the adoption of digital pedagogy in university settings.

2. Literature Review

The digitalization of education has created new opportunities for teaching and learning (Pillai et al., 2023; Savotina et al., 2020). However, many educators have yet to adopt digital pedagogy fully. They need to change their mindsets and redefine their roles (Herbert et al., 2020). Digital pedagogy (DP) involves using digital tools thoughtfully within an educational framework (Khan, 2021). It combines both constructivist and traditional teaching methods. Traditional approaches focus on teacher-led instruction, while the constructivist model emphasizes student-centered learning, collaboration, and active engagement (Väätäjä & Ruokamo, 2021). Research shows that constructivist methods support technology integration in classrooms more effectively (Pittman & Gaines, 2015). Another important framework for digital pedagogy is TPACK. This framework helps educators design meaningful and engaging learning experiences. Studies indicate that teachers' TPACK skills are crucial for implementing digital strategies successfully. This reinforces the need for ongoing professional development (Souza & Cardoso, 2024; Rosamsi & Nurdiani, 2024; Silvester et al., 2024). Theories such as Innovation Diffusion Theory (Rogers, 1962), Technology Acceptance Model (Davis, 1989), and UTAUT (Venkatesh et al., 2003) also shed light on what affects technology adoption in education. Despite the increasing research on technology adoption in higher education, there are still few studies focused on digital pedagogy (Belenkova, Skudnyakova & Bosov, 2022). In India, the adoption of digital pedagogy has increased, particularly due to COVID-19 and government initiatives. Programs aimed at improving digital infrastructure and skills have expanded access to online courses and virtual classrooms, especially in remote areas (Kumar, 2024; Singh, 2023; Fatima et al., 2025). However, challenges such as digital equity, inadequate teacher training, and limited internet access in rural areas continue (Roy, 2022; Singh, 2023). These issues raise concerns about educational quality and the digital divide, particularly for disadvantaged learners.

Earlier studies on digital pedagogy adoption have largely focused on primary and secondary education teachers in advanced economies, identifying both institutional barriers (limited access to resources, training, and support) and personal barriers related to teachers' confidence, beliefs, attitudes, and perceived value of technology (Ertmer et al. 2012; Pongsakdi, Kortelainen & Veermans, 2021). Similarly, teachers' attitudes towards technology have been identified as an important parameter that may significantly impact their perception of the usefulness of technology (Instefjord and Munthe 2017; Teo, Zhou & Noyes, 2016). Further, El-Hamamsy et al. (2024) have also identified the challenges encountered by the school teachers in the adoption of digital pedagogy, for instance, lack of access to resources, fear of change, workload increase, and the need for high-quality digital learning materials. Another study in the Australian and Swedish context emphasized competencies such as attitude, self-efficacy, and peer collaboration skills that can contribute to the successful integration of technologies in teaching (McCarthy, Maor & McConney, 2017; Mannila, 2018).

So far as the Indian educational landscape is concerned, the research studies here have primarily emphasized the quantitative exploration. To quote a few research studies, Sharma and Srivastava (2020) empirically confirmed the significant positive impact of value beliefs (VB), social influence (SI), and perceived ease of use (PEOU) on the behavioural intention (BI) to use technology by the teachers in management institutions. Additionally, another research study in Indian HEIs examined the perceptions of teachers and found that continuance intention towards using online teaching in HEIs is most significantly influenced by teachers' satisfaction rather than perceived usefulness (PU), perceived ease of use (PEOU), and attitude (Kumar et al., 2022).

Further, only a limited number of qualitative studies have been conducted. For instance, Goarty and Gupta (2023) examined the factors influencing digital transformation in Indian higher education institutions and identified environmental factors and teachers' knowledge as key predictors of digital transformation. Phutela and Dwivedi (2020), using IPA, identified inhibitors and motivators of e-learning adoption by students. Another qualitative exploratory case study focused on the Indian elementary school teachers and unveiled their experiences of the learner-centered pedagogy (Evans, 2023). The available literature reveals a dearth of studies examining teachers' perspectives (Singh, Sharma & Paliwal, 2021). Most existing studies have primarily employed the Technology Acceptance Model and are empirical in nature. The majority of this research has focused on students' digital learning acceptance behaviour rather than on educators' experiences. Furthermore, the literature indicates a lack of qualitative exploration of technology integration within the teaching–learning process (Kaushik and Verma, 2020). Therefore, this knowledge and methodological gap paves the way for exploring teachers' perspectives on digital pedagogy adoption using the Interpretative Phenomenological Analysis approach within Indian higher education institutions (HEIs).

3. Research Methodology

3.1 Research Design

IPA has been used in the current study to identify the factors affecting the digital pedagogy adoption amongst teachers in Indian HEIs. This technique is usually applied to a research question or problem that needs to be discovered and is at a nascent stage, where not much research has been conducted. Further, this technique has found its popularity in the field of education research as well (Bhaskar & Rana, 2024; Rana, 2022; McCarthy, Glassburn & Dennis, 2022; Creely & Laletas, 2020). In the current study, the authors unveil the factors affecting the University teachers' adoption of digital pedagogy. IPA allowed the authors to conduct in-depth interviews and bring out the ground perceptions of the University faculty members about the challenges they face while implementing digital pedagogy in their daily routine.

Following the approach of Eatough and Smith (2017), the IPA technique was incorporated into this study for its idiographic focus, holistic orientation, and phenomenological depth. The authors followed the guidelines of Bogner, Littig & Menz (2009) and Patton (1990), where semi-structured interviews were conducted with homogeneous respondents which consisted of faculty members in the higher education sector. In-depth analysis of the existing literature provided the authors with clarity on basic terms like digital pedagogies and the already studied factors in this and related research topics. This helped to build the initial draft of the interview questions, which were verified by a group of academic experts who dealt in quality assurance in different Universities and faculty members with different experiences in teaching.

3.2 Respondent Profile

The respondents consisted of faculty members in different Universities and colleges. The authors deliberated to include teaching faculty members from different parts of the country and also from different tier cities to gain as diverse inputs as possible. An effort was made to include Faculty from diverse fields of expertise to understand if there was any difference in the same in adoption of digital pedagogies. The sample consisted of 44% male and 56% female Faculty with an average work experience of 8 years (Refer to Table 1)

Table 1: Respondents' Profile

Respondent	Domain	Designation	Location of the Institution
R1	Engineering	Associate Professor	Gurgaon
R2	Finance	Assistant Professor	Dehradun
R3	Management	Professor	Chandigarh
R4	Design	Associate Professor	Mumbai
R5	Law	Assistant Professor	Vadodara
R6	Marketing	Assistant Professor	Bangalore
R7	Media	Assistant Professor	Bhuwaneshwar
R8	Liberal Studies/ Humanities	Assistant Professor	Noida
R9	Data Analytics	Associate Professor	Delhi
R10	Taxation	Assistant Professor	Calcutta
R11	Engineering	Assistant Professor	Mandi
R12	Human Resource	Assistant Professor	Rishikesh
R13	Media	Professor	Dehradun
R14	Media	Assistant Professor	Delhi
R15	Health Sciences	Assistant Professor	Rishikesh
R16	Design	Assistant Professor	Delhi
R17	Finance	Professor	Indore
R18	Law	Assistant Professor	Haridwar

3.3 Data Collection

The data collection phase spanned from February to March 2024, during which semi-structured interviews were conducted with 18 purposively selected faculty members. The sampling strategy followed guidelines for Interpretative Phenomenological Analysis (IPA), where information-rich cases are prioritized (Palinkas et al., 2015; Creswell & Creswell, 2017).

Respondents were identified based on two key inclusion criteria:

- Active or prior implementation of digital pedagogical tools in their teaching; and
- Diversity in academic disciplines and geographical locations (urban vs. semi-urban/tier cities).

LinkedIn served a dual purpose: it was used both to identify suitable participants, based on their public teaching portfolios, and to establish initial contact. This approach ensured a homogenous sample in terms of exposure to digital pedagogy, consistent with the idiographic emphasis of IPA (Smith, Jarman & Osborn, 1999).

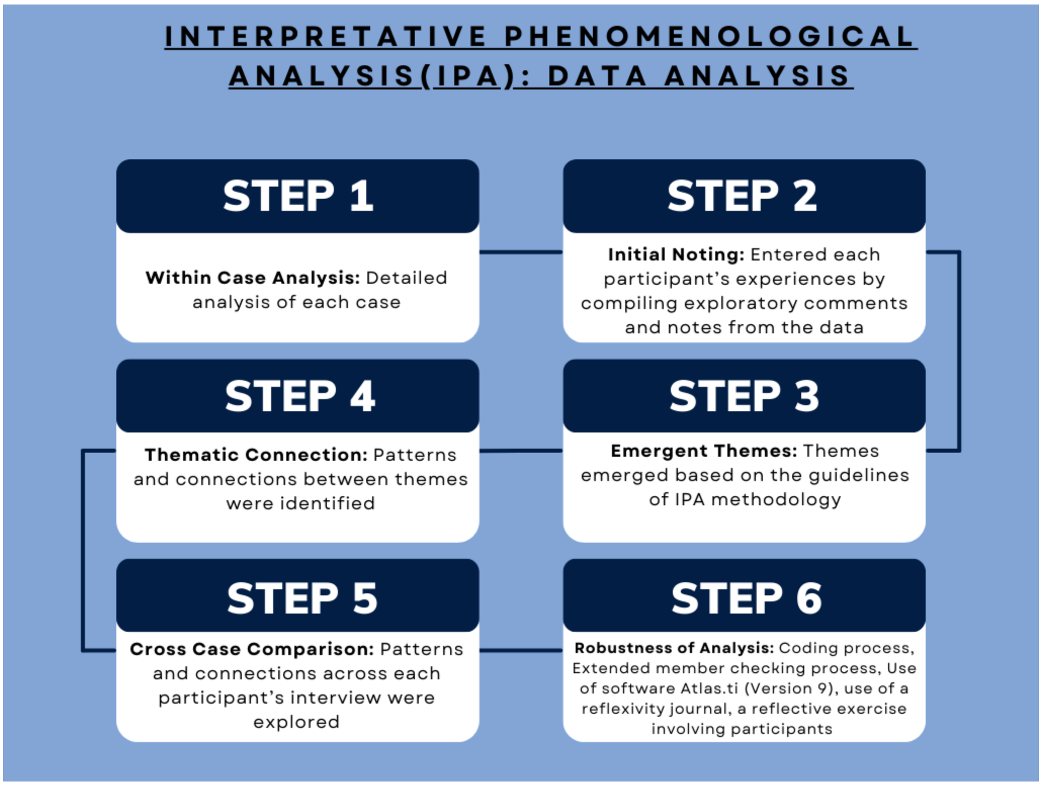
The final number of respondents (n = 18) was not predetermined but determined through data saturation. After the 16th interview, no new sub-themes emerged, indicating thematic redundancy. Two additional interviews were conducted to confirm this saturation point (Rajasinghe, 2020; Vasileiou et al., 2018).

All interviews were conducted remotely via Zoom or telephone to ensure the convenience and availability of the respondents. Each session lasted 35–45 minutes, a duration that aligns with best practices in IPA for eliciting rich, reflective narratives while avoiding fatigue (Eatough & Smith, 2017; Peat, Rodriguez & Smith, 2019; Fan et al., 2024; Bauman, 2015).

Participants were informed about the study's objectives and terminology prior to their interviews. Informed consent was obtained for audio recording. For those who declined recording (three respondents), detailed field notes were taken and verified post-interview.

3.4 Data Analysis

Interpretative Phenomenological Analysis (IPA) was employed to guide the data analysis process in this study (refer to Figure 1). Rooted in its theoretical and philosophical underpinnings, the authors adopted IPA's analytical framework (Smith, Jarman & Osborn, 1999; Alase, 2017; Larkin, Shaw & Flowers, 2019), which involved verbatim coding of raw data, interpretative engagement, and theme development (Nizza, Farr & Smith, 2021). To enhance analytical rigour, the authors employed *auditing techniques*, which in IPA refer to a systematic approach to ensure transparency and consistency in coding and interpretation by maintaining an audit trail of decisions made throughout the analysis (Smith, Jarman & Osborn, 1999). Each author independently analysed the data and later engaged in collaborative discussions. This comparative approach facilitated an extended member checking process and contributed to the credibility of the findings. Themes and sub-themes showed high concordance across individual analyses, a consistency that was confirmed by the use of Atlas.ti (Version 9), which showed approximately 94% agreement.



Source: Smith & Osborn (2008); Moustakas (1994)

Figure 1: Interpretative phenomenological analysis (IPA)-Data Analysis

To address personal biases, a *reflexivity journal* was maintained by the researchers, especially during the initial interviews. After each of these early interviews, authors spent approximately twenty minutes writing down reflective notes, observations, and potential thematic connections (Peat, Rodriguez & Smith, 2019). This reflexive journaling supported ongoing critical reflection, bracketing of assumptions, and enhanced the confirmability of the study (Vicary, Young & Hicks, 2017; Engward & Goldspink, 2020; Goldspink & Engward, 2019).

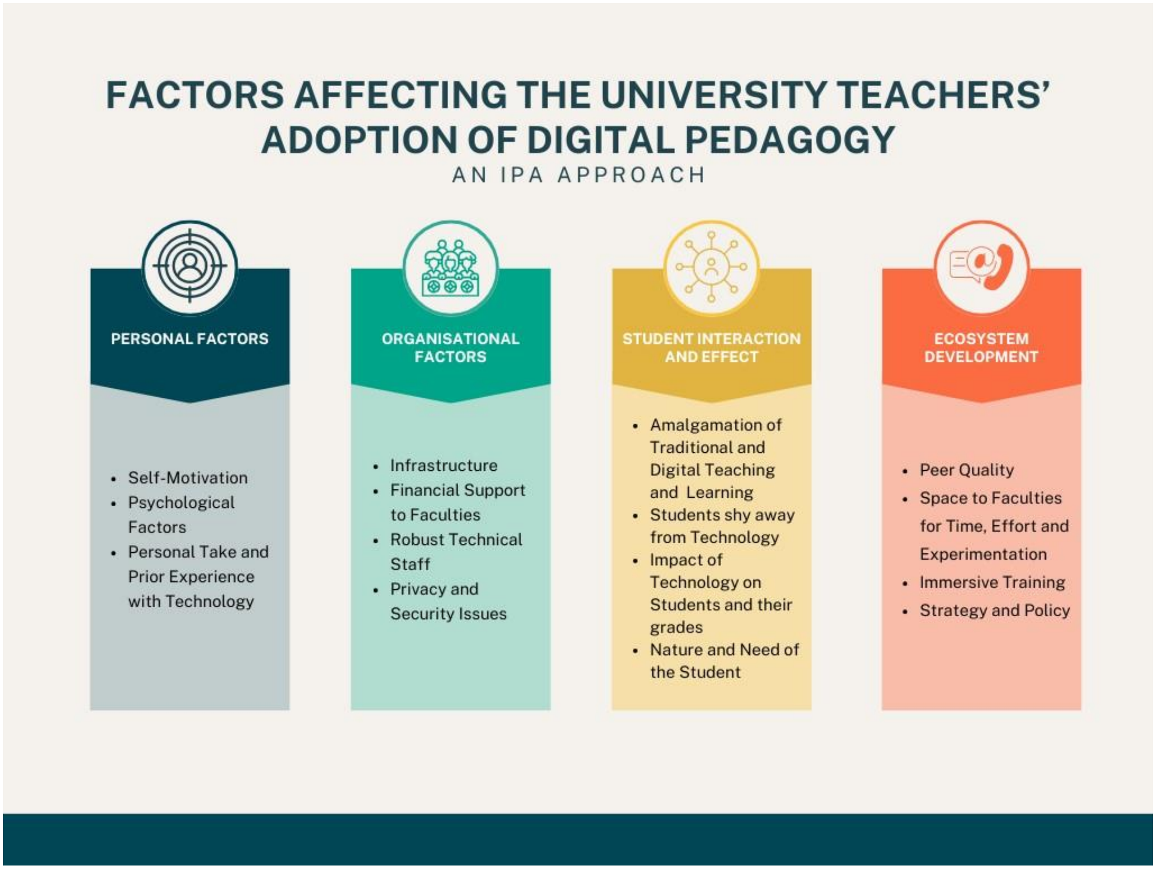
The researchers also applied the *hermeneutic circle*, a core IPA concept referring to the iterative process of understanding parts of the data in relation to the whole, and vice versa, to interpret participants' lived experiences within their broader contexts.

A reflective exercise was conducted during a second round of interviews, wherein participants were presented with preliminary findings derived from the first round. This process served as a member checking mechanism to validate the emerging interpretations (Smith & McGannon, 2018). All participants confirmed that the interpretations accurately represented their experiences, thereby enhancing the credibility of the findings.

Out of the total interviews conducted, three were not audio-recorded due to participant preference. In these instances, the researchers relied on detailed note-taking during the interview and immediately afterwards. For member checking in these cases, the researchers shared synthesized summaries of interpreted themes with the respective participants through follow-up communication, allowing them to review, validate, and revise any part of the interpretation as needed. No discrepancies were noted during this validation process.

4. Results and Findings

The findings of the study revealed four major themes, namely, personal factors, organisational factors, student interaction and effect, and ecosystem development. The themes and sub-themes are discussed in length under this section (refer to Figure 2). The discussion addresses each of the three research questions guiding this study: (Q1) teachers' lived experiences with adopting digital pedagogy, (Q2) key factors influencing adoption, and (Q3) strategies that facilitate effective implementation within Indian HEIs.



Source: Authors' own work

Figure 2: Factors Affecting the University Teachers' Adoption of Digital Pedagogy

Theme 1: Personal Factors

In response to Research Question 1 (Q1)—How do Indian university teachers make sense of their lived experiences with adopting digital pedagogy in higher education?—the first theme that emerged was titled as 'Personal Factors', which included the following sub-themes (refer to Table 2)- 'Self-Motivation', 'Psychological factors', 'Personal take and Prior Experience with Technology'. The elaboration of each sub-theme, along with the responses, has been explained below:

Table 2: Sub-themes, Responses, and Key Takeaways- Theme 1: Personal Factors

Theme 1: Personal Factors	
Sub-theme 1: Self-Motivation	
Responses	**Key Take-Aways**
R5: "It is self-motivation that drives an individual to learn more on their own. Imbibing technology is possible only through a self-driven faculty."	Self-driven Faculty are technology inclined.
R3: "Faculty do the bare minimum when it comes to using technology. It is because of the huge workload, leaving little time to explore and experiment with the existing digital tools."	Faculty find it difficult to use technology in their teaching pedagogy because of the heavy workload.
R6: "I teach a theory-based course where technology isn't essential; digital tools are more relevant in subjects requiring practical exposure."	Faculty finds relevance of digital tools in practical courses rather than theory courses.
R14: "Learning technology is entirely a teacher's personal effort as there is no mandate to use the digital tools in the class."	There is no mandate to use digital tools in classroom teaching.
Sub Theme-2 Psychological Factors	
Responses	**Key Take-Aways**
R8: "I prefer traditional teaching, as digital tools tend to make students lazy and passive. They follow instructions but struggle to apply concepts independently."	More comfortable with the traditional teaching method. The use of digital tools while teaching makes students lazy and inactive in class.
R12: "I am not very comfortable with using digital pedagogical tools as there is always a possibility of making more errors."	Faculty members believe that there is a possibility of more errors while using digital tools.
R2: "I avoid experimenting with digital tools as they require time and effort to learn, and are not always welcomed by students or the institution."	Faculty have discouraging experiences while using digital tools.
Sub-theme 3: Personal Take and Prior Experience with Technology	
Responses	**Key Take-Aways**
R13: "I have a background in computer science; therefore, I am comfortable with using and experimenting with different digital tools for teaching."	Prior experience with technology helps in inculcating digital tools in teaching
R1: "I am a strong believer in the amalgamation of technology with education. I started experimenting with different digital tools I learnt during different FDPs (Faculty Development Program) and MDPs. (Management Development Program)"	Personal take on technology influences the use of digital tools in teaching.

The sub-theme *'Self-Motivation'* highlights its vital role in integrating technology into teaching. Faculty with intrinsic motivation proactively explore and adopt digital tools, regardless of institutional mandates (Balakrishnan & Shuib, 2021; Rahi et al., 2021). While these self-driven educators embrace innovation, factors like heavy workloads and course-specific demands can hinder adoption.

The sub-theme *'Psychological Factors'* captures the cognitive and emotional challenges faculty face in adopting digital pedagogy. Use of new technologies often led to more frequent errors—such as typos, uploading issues, and execution mistakes—compared to traditional methods (Gerli et al., 2022). Negative classroom experiences and student feedback further fuelled doubts about technology's effectiveness. Moreover, faculty's comfort with conventional methods, coupled with perceived low student engagement in digital settings, significantly deterred adoption (Ray, Bala & Dasgupta, 2019).

The sub-theme *'Personal Perspective and Previous Encounters with Technology'* reveals that faculty with computing proficiency adopt digital tools more easily and are open to experimentation. Those with positive prior experiences actively pursue refresher courses, showing a strong willingness to enhance efficiency through technology (Tursunbayeva & Gal, 2024).

Theme 2: Organisational Factors

In continuing to the discussion aligned with Research Question 2 (Q2)—*What are the key factors influencing university teachers' adoption of digital pedagogy in Indian higher education?*—the theme of 'Organisational Factors' (refer to Table 3)- and includes the following sub-themes- 'Infrastructure', 'Financial Support to Faculty', 'Robust Technical Staff', and 'Privacy and Security Issues'. This section enumerates each sub-theme by laying down the following responses and the key takeaways.

Table 3: Sub-themes, Responses, and Key Takeaways- Theme 2: Organisational Factors

Theme 2: Organisational Factors	
Sub-theme 1: Infrastructure	
Responses	**Key Take-Aways**
R3: "There is a need for huge capital investment to build the required digital infrastructure. This will help the faculty to use the digital pedagogies."	Infrastructure building is required for making use of advanced digital tools and technology. Infrastructure building requires huge capital investment.
R17: "We are still struggling with the basic infrastructure, like internet connectivity. Example out of 60 systems, only 10-15 systems will be working efficiently."	Basic infrastructure still requires overhauling to achieve 100% efficiency.
R18: "We are struggling to get subscriptions for basic statistical software, plagiarism check software, journal subscriptions, and availability of smart classes."	Basic digital infrastructure and necessities to become an effective faculty are still required to be fulfilled.
Sub-theme 2: Financial Support to the Faculty	
Responses	**Key Take-Aways**
R17: "I wanted to learn the inclusion of simulation in my domain, which cost me around 50,000 rupees, as no financial support was provided to me by the institution. Also, my leaves were used during travelling."	Lack of financial support to learn advanced digital tools in teaching
R15: "To learn technology for digital pedagogy requires some financial support from the institution, which is still missing."	Revamping policy for financial support for faculty training and development.
Sub-theme 3: Robust Technical Support	
Responses	**Key Take-Aways**
R11: "Once I had prepared all the study material with the help of digital tools, but was not able to deliver it in the class due to technical issues, so I switched to board and chalk teaching."	Frequent requirement for technical staff
Sub-theme 4: Privacy and Security Issues	
Responses	**Key Take-Aways**
R4: "I learnt to use a digital tool, but avoid it as the institution requires us to teach using personal systems. Logging in with email-linked passwords and connecting to multiple devices raises privacy concerns."	Using own systems for Institutional work creates privacy issues, especially where login is connected to their personal email IDs.
R7: "I have experienced my personal system getting corrupted when connecting with the institution's devices. This has resulted in the loss of my personal research work as well."	Connecting personal systems with different devices causes a loss of personal data of the faculty.
R13: "It's unsafe to use personal systems for institutional work, risking data leaks. But due to limited funding, faculty have no option but to rely on their own devices."	Faculty working on their systems can lead to leakage of sensitive institutional data.

The primary sub-theme highlights the financial commitment institutions must make to build the digital *infrastructure* essential for adopting digital pedagogy (Mateko, 2024). Faculty emphasized that integrating advanced digital tools requires robust infrastructure, which motivates their engagement with these platforms. They noted that investing in updated technologies helps them stay current and experiment with innovations. Currently, many faculty feel the existing institutional technologies are inadequate and call for a comprehensive upgrade (Verdecchia, Lago & Vries, 2022).

The second sub-theme, *'Financial Assistance for Faculty Members,'* reveals faculty concerns over insufficient institutional support for learning new digital technologies (Almansour & Almoayad, 2024). Faculty faced difficulties securing leave and funding to attend training at prestigious institutions. Existing policies often exclude travel and accommodation costs, forcing them to bear expenses personally or miss these opportunities entirely.

The third sub-theme- ' Robust *Technical Support'* highlights the necessity for a proficient technical team to assist faculty members in setting up classroom technology and addressing any mid-session technical challenges promptly (Kumalasari et al., 2024). Faculty members encounter various instances where inadequately trained technical staff leads to disruptions in lecture delivery, resulting in valuable time loss for both faculty members and students. This aspect led to the non-adoption of digital pedagogy, sticking to traditional methods of teaching.

The fourth sub-theme, *'Security and Privacy Issues,'* highlights faculty concerns over using personal devices connected to institutional networks (Kumar et al., 2022). Many reported data and research losses due to unsecured networks, leading to distrust and reluctance toward digital pedagogy. The gap between policy and practice, along with inadequate cybersecurity support, exacerbates these fears. Institutions must offer secure infrastructure, technical assistance, and clear data protection policies to build faculty confidence and enable sustainable digital adoption.

Theme 3: Student Interaction and Effect

Also connected to Research Question 2, the third theme, *Student Interaction and Effect (Refer to Table 4)*, expands the understanding of influencing factors by focusing on how students' engagement, attitudes, and learning outcomes shape faculty adoption decisions and their perception of the value of digital pedagogy. The third theme includes the following sub-themes: 'Amalgamation of Traditional and Digital Teaching and Learning', 'Students Shy Away from Technology', 'Impact of Technology on Students and their Grades' and 'Nature and Needs of Students'. This section enumerates each sub-theme by laying down the responses of faculty members and the key takeaways.

Table 4: Sub-themes, Responses and Key Takeaways-Theme 3: Student Interaction and Effect

Theme 3: Student Interaction and Effect	
Sub-theme 1: Amalgamation of Traditional and Digital Teaching and Learning	
Responses	**Key Take-Aways**
R11: "Relying only on technology for teaching leaves the students with no interest in the faculty. It is the one-on-one interaction between students and faculty that makes learning most effective."	Digital tools are supplementary to traditional teaching. Traditional teaching is irreplaceable. Amalgamation of traditional and digital teaching is the most effective.
F17: "I use digital pedagogical tools but continuously engage students with questions, activities, opinions, and numerical problems during class to maintain their involvement."	
R5: "In face-to-face teaching, factors like body language and real-time modulation allow customization based on students' needs, which isn't possible in online mode."	
Sub-theme 2: Students Shy Away from Technology	
Responses	**Key Take-Aways**
R10: "Many students shy away from using technology for learning, struggling with software and doing only the minimum to show task completion."	Students shy away from using digital tools. Students are not using the available digital resources to their advantage.

Theme 3: Student Interaction and Effect	
R16: "Only a few students actively use technology for studying, while most underutilize available digital resources. To address this, we've been conducting awareness workshops at our institution."	Awareness workshops are a way to create the required push in effective digital learning.
Sub-theme 3: Impact of Technology on Students and their Grades	
Responses	**Key Take-Aways**
R1: "I've used digital tools for years, but student grades haven't improved significantly; they focus only on exam prep, not on learning new concepts."	There is no radical difference in the grades of the students after including digital tools for teaching.
R3: "I've integrated advanced digital tools in teaching, but only 5–10% of students respond well; most find them daunting or are reluctant to learn quickly."	A very low percentage of students learn from the advanced digital tools.
Sub-theme 4: Nature and Needs of Students	
Responses	**Key Take-Aways**
R15: "Digital tools help students with language barriers by providing accessible content they can revisit to clarify doubts and solve problems."	Digital tools and platforms have provided students with flexibility in relation to access to study material and content.
R1: "Digital pedagogy has helped us identify slow and fast learners, enabling us to tailor supportive materials and activities using advanced tools to meet their diverse needs."	Digital tools help in customising learning requirements as per students' calibre.
R14: "Recent curriculum changes have added vocational courses focused on new technologies to enhance student employability."	Digital capabilities are a mandate to enhance student employability.

The first sub-theme, '*Amalgamation of Traditional and Digital Teaching and Learning,*' reflects faculty views that face-to-face interaction surpasses digital delivery (Schmitz et al., 2024). Faculty strongly believe that combining traditional and digital methods achieves optimal efficiency and student outcomes (Harper, McCormick & Marron, 2024). This blended approach offers flexibility, enabling faculty to tailor sessions for the best learning results.

The second sub-theme, *'Student Shy Away from Technology,'* highlights faculty concerns about students' reluctance to embrace digital tools, varying by background and city tier (Rahimi et al., 2024). Many undergraduates and postgraduates lack awareness of digital resources offered by HEIs, limiting their effective use. This can be addressed through proactive awareness programs by HEI libraries and digital teams. Tracking library visits and resource usage helps institutions identify gaps and tailor interventions.

The third sub-theme, *'Impact of Technology on Students and their Grades,'* reflects faculty observations of no significant grade improvement after integrating digital tools. They also noted increased student reliance on these platforms to complete tasks, reducing effort (Harper, McCormick & Marron, 2024; Osabutey, Senyo & Bempong, 2024).

The fourth sub-theme, *'Nature and Needs of Students,'* highlights how digital tools help faculty identify diverse learning paces and tailor support accordingly (Hernandez & Keane, 2024). These platforms provide access to teaching materials and personalized resources, while vocational courses on sites like Coursera and Udemy enhance employability and enable customized advanced courses.

Theme 4: Ecosystem Development

Addressing Research Question 3—*What strategies facilitate the effective adoption of digital pedagogy in HEIs?*—the final theme, *Ecosystem Development* (refer to Table 5), identifies actionable strategies at both institutional and systemic levels. These include peer mentorship programs, immersive training, flexible experimentation spaces, and policy reforms that collectively promote sustained and meaningful digital adoption. This section enumerates each sub-theme by laying down the Responses and the key takeaways.

Table 5: Sub-themes, Responses and Key Takeaways-Theme 4: Ecosystem Development

Theme 4: Ecosystem Development	
Sub-theme 1: Peer Quality	
Responses	**Key Take-Aways**
R15: “Research competition is intense, with little peer collaboration. Experienced faculty rarely assist with new digital research and teaching tools, leaving individuals to invest their own time, money, and effort despite having in-house experts.”	Poor peer collaboration and support
R7: “Among my peers, I use digital teaching tools the most. The environment isn’t competitive, and most faculty prefer traditional methods since digital tools aren’t mandatory.”	Less interest of peers in discovering new digital tools for teaching.
Sub-theme 2: Space to Faculty for Time, Effort and Experimentation	
Responses	**Key Take-Aways**
R18: “We are often pressed against time to learn new things and enhance our skill sets. To deliver a session via digital modes requires a lot of background time.”	Use of digital tools requires considerable background work.
R4: “Faculty need proper training, time, and support to integrate technology effectively; mandating digital tools risks only surface-level, minimal use.”	Mandating digital tools will only help inculcate the same on a very surface level.
R11: “Before COVID, our institution introduced an LMS, starting with basic features due to initial challenges. Gradually, both faculty and students adapted, allowing us to later implement advanced features that eased our work during the pandemic.”	Use of digital tools takes time to imbibe.
R9: “The effort faculty invest varies individually. They need time and freedom to experiment with digital tools, accepting mistakes as part of the learning and adaptation process.”	Effective use of digital tools requires flexibility to experiment.
Sub-theme 3: Immersive Training	
Responses	**Key Take-Aways**
R16: “The organization can ease psychological barriers by providing extensive, hands-on training that lets faculty immerse themselves, experiment, and gain confidence with digital tools.”	Hands-on immersive training can help remove psychological barriers.
R13: “Learning digital tools and platforms requires time and ongoing support; a few online sessions aren’t enough for daily adoption.”	Imbibing technology will take time for the faculty.
Sub-theme 4: Strategy and Policy	
Responses	**Key Take-Aways**
R10: “The training should be need-based after conducting the need analysis with the help of fellow faculty.”	Requirement of need-based training sessions. Experts should be selected after the need analysis is conducted.
R5: “Integrating advanced digital tools requires a clear strategy; faculty shouldn’t be overwhelmed but gradually accustomed, allowing them to become proficient at their own pace.”	Use of advanced digital tools in teaching pedagogy requires a proper strategy in place.
R14: “Institutions can really assist the Faculty in building digital competencies by revamping the financial and training support strategy and policies.”	HEIs require revamping the financial and training support strategy and policies.

The first sub-theme, *‘Peer Quality,’ highlights colleagues’* strong influence on adopting digital pedagogy. Faculty noted minimal peer engagement with digital tools, shaping group behaviour (Kim et al., 2025). Institutional initiatives driven by senior and motivated junior faculty, supported by management's commitment to quality and flexibility, foster consistent development and integration of new technologies across the HEI.

The second sub-theme, *'Space for Time, Effort, and Experimentation,'* highlights that faculty need ample time and ongoing engagement with technology to become comfortable with digital delivery, despite mastering course content (Laurillard, 2024). Institutions must provide this space, identifying motivated faculty to collaborate with digital teams on tailored solutions.

The third sub-theme, *'Immersive Training,'* stresses the need for continuous, in-depth training in advanced technologies like simulations, AI, and the metaverse (Negahban, 2024). Such comprehensive training, requiring significant HEI investment, enables faculty to confidently and independently use these tools, meeting high institutional standards.

The fourth sub-theme, *'Strategy and Policy,'* explores how HEIs can tailor strategies to promote faculty adoption of digital pedagogy (Rana et al., 2023; Huda, 2019). This includes conducting needs analyses, selecting expert trainers, and involving quality assurance experts in committees to enable faculty visits to reputable institutions and stay updated on emerging technologies. Moreover, HEI policies on travel, accommodation, and financial support must be restructured to better facilitate faculty training and digital skill development.

In response to the above section, the study identifies specific strategies that facilitate digital pedagogy adoption in higher education institutions (HEIs), particularly in the Indian context, where infrastructural, psychological, and policy-related challenges often impede integration. The following strategies emerged from the thematic analysis of faculty interviews:

1. Institutionalizing Peer Mentorship Programs: One of the significant insights from the study is the lack of collaborative culture among faculty members when it comes to digital pedagogy. Many digitally adept faculty members work in isolation, while those less confident remain hesitant to experiment due to a lack of guidance. To bridge this gap, HEIs should develop structured mentorship programs where technologically proficient faculty are paired with peers who require support. These mentoring relationships can include joint classroom sessions, peer reviews of digital content, or informal 'tech clinics' where faculty troubleshoot digital tools together. Such peer-to-peer engagement not only facilitates skill transfer but also helps build trust and reduce anxiety associated with digital teaching.
2. Embedding Financial Support into Institutional Policy: The financial burden of learning new digital tools and attending training programs was consistently cited by respondents as a barrier. Many faculty members expressed that they had to self-fund their professional development due to inadequate institutional support. To address this, HEIs should revise their faculty development policies to include dedicated funding for technology-related upskilling. This could cover costs for attending national and international workshops, subscribing to professional platforms (such as Coursera or LinkedIn Learning), or accessing licensed simulation tools. Additionally, granting academic leave without penalizing faculty workloads for such initiatives would demonstrate institutional commitment to digital transformation.
3. Designing Need-Based, Customized Training Modules: The study also highlights dissatisfaction with generic training programs that do not cater to individual skill levels or discipline-specific needs. Faculty with varying degrees of digital proficiency benefit from different types of support. For instance, a novice may need foundational training on Learning Management Systems, while an advanced user may seek exposure to immersive tools like AR(Augmented Reality)/VR (Virtual Reality) or AI (Artificial Intelligence)-based assessment platforms. Therefore, HEIs should begin with a diagnostic needs assessment to identify existing skill gaps and then curate modular training sessions accordingly. This targeted approach will ensure that faculty feel the training is relevant, manageable, and aligned with their teaching responsibilities.
4. Providing Time and Space for Digital Experimentation: A recurring theme in the interviews was the lack of time and institutional encouragement for experimentation. Faculty often reported that the pressure of completing syllabi and administrative tasks leaves little room for testing new digital tools. Without the freedom to explore, innovate, and even fail occasionally, faculty are unlikely to move beyond surface-level adoption. Institutions must therefore allocate 'protected time'—through reduced teaching loads or dedicated innovation hours—so faculty can meaningfully integrate digital tools into their pedagogy. Additionally, setting up sandbox environments or digital learning labs where faculty can test technologies without consequences will promote confidence and sustained usage.

Taken together, the findings across all four themes provide a holistic response to the three research questions. They reveal that while individual agency (Q1) and organizational context (Q2) play crucial roles in shaping adoption, long-term success depends on ecosystem-level strategies (Q3) that institutionalize support and foster a culture of innovation.

5. Implications

5.1 Theoretical Implications

This research paper contributes significantly to the theoretical landscape of digital pedagogy adoption within Higher Education Institutions (HEIs). By offering in-depth qualitative insights into the use of digital pedagogy by faculty members, the study enriches the existing literature on digital adoption. While theories such as the Technology Acceptance Model (TAM), Unified Theory of Acceptance and Use of Technology (UTAUT), and diffusion of innovation theory have predominantly been explored in quantitative contexts, this research breaks new ground by providing a qualitative analysis that delves into the psychological aspects of the usage and adoption of digital pedagogy by teachers in HEIs.

Furthermore, the study addresses contemporary topics such as digital pedagogy, quality education delivery, enhancing teaching pedagogy via digital tools, and faculty development, all of which have gained significant prominence in the academic field due to their practical implications. The emergence of higher education quality norms, exemplified by initiatives like the National Education Policy (NEP) and accreditation and ranking bodies, has brought these issues to the forefront of academic discourse.

5.2 Practical Implications

The study has the following practical implications:

Faculty Members: The study offers valuable practical implications for faculty members. Firstly, it enhances faculty members' awareness by providing insights into the prevailing perceptions surrounding digital pedagogy, enabling them to navigate the existing norms more effectively. Second, the study offers specific solutions to common problems encountered in the use of digital tools for teaching and learning. Third, the insights from the study foster collaboration among faculty members, promoting peer support and continuous learning in digital pedagogy.

HEIs Management: Management can identify areas where policy and infrastructural changes are needed to support the integration of digital tools into daily teaching practices. This understanding enables management to enact necessary reforms and provide the requisite resources to facilitate the effective implementation of digital pedagogy across the institution. HEIs can create an environment conducive to the smooth incorporation of digital tools into teaching methodologies, ultimately enhancing the quality of education and student learning outcomes.

Quality Assurance Experts: The study provides insights for quality assurance experts within Higher Education Institutions (HEIs) as well as freelance experts. By understanding the ground issues and challenges faced by faculty members in implementing digital tools in teaching and learning pedagogy, these experts can provide valuable guidance and support. The research elucidates both the resource-related challenges, such as access to technology and training, as well as the psychological factors, such as resistance to change and fear of technology. This could involve developing tailored training programs, providing access to necessary resources, and implementing support systems to assist faculty members in overcoming barriers to digital integration.

Policymakers: The academic research paper offers practical implications that extend to policymakers as well. The understanding of challenges in the adoption of digital pedagogy serves as a foundation for developing informed policies aimed at fostering a more conducive environment for digital pedagogy by developing more tailored and effective faculty refresher courses. By addressing faculty members' concerns and reservations about technology use, training programs can be designed to be more hands-on and practical, thereby enhancing faculty members' digital literacy and confidence in incorporating digital tools into their teaching practices.

6. Conclusion

In conclusion, this study offers valuable insights into the multifaceted factors influencing university teachers' adoption of digital pedagogy within Indian higher education institutions, thereby addressing RQ1, which examined the institutional, technological, and personal factors shaping digital pedagogy adoption. Using Interpretative Phenomenological Analysis, and by foregrounding the lived experiences of faculty across diverse domains, the study also addresses RQ2 by elucidating how teachers perceive, experience, and make sense of digital pedagogy in their teaching practices. Together, the findings presented in Section 4 and synthesized through the implications discussed in Section 5 highlight the interdependence of individual agency, institutional support, and systemic conditions in enabling meaningful digital integration.

However, several limitations must be acknowledged. First, the qualitative design and purposive sample of 18 university faculty members, while enabling in depth insight, may limit generalizability across the diverse higher education landscape in India. Second, reliance on self reported interview data may introduce subjective bias, despite measures such as reflexivity and member checking. Third, the study focuses primarily on faculty perspectives and does not incorporate views of other stakeholders such as students, administrators, or digital infrastructure teams, whose inclusion could enrich understanding of digital pedagogy adoption.

Future research may extend the findings related to RQ1 through larger scale, mixed method, or quantitative studies to validate and generalize the identified enabling and constraining factors. Similarly, further exploration of RQ2 through comparative or cross institutional studies may deepen understanding of how contextual and policy environments shape teachers' digital pedagogical experiences. Investigating emerging technologies such as artificial intelligence, immersive learning environments, and adaptive systems also represents a promising direction for future inquiry.

Ultimately, strengthening alignment between institutional strategies and teachers' lived experiences, as highlighted across Sections 4, 5, and 6, will be critical for designing inclusive, sustainable, and resilient digital transformation pathways in higher education.

AI Statement: The authors declare that AI tools were not used in this study for the paper's conception, revision, or the creation of figures and tables.

Ethical Approval Statement: This study was conducted in accordance with the Academic Integrity Code

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Competing Interests: The authors declare that there is no conflict of interest.

References

Alase, A., 2017. The interpretative phenomenological analysis (IPA): A guide to a good qualitative research approach. International journal of education and literacy studies, 5(2), pp.9-19.

Alhubaishy, A. and Aljuhani, A., 2021. The challenges of instructors' and students' attitudes in digital transformation: A case study of Saudi Universities. Education and Information Technologies, 26(4), pp.4647-4662.

Almansour, M. and Almoayad, F., 2024. Exploring challenges and perceptions in the learning environment: an online qualitative study of medical students. BMC medical education, 24(1), p.147.

Balakrishnan, V. and Shuib, N.L.M., 2021. Drivers and inhibitors for digital payment adoption using the Cashless Society Readiness-Adoption model in Malaysia. Technology in Society, 65, p.101554.

Bali, S. and Liu, M.C., 2018, November. Students' perceptions toward online learning and face-to-face learning courses. In Journal of Physics: conference series (Vol. 1108, p. 012094). IOP Publishing.

Bauman, A., 2015. Qualitative online interviews: Strategies, design, and skills. Qualitative research in organizations and management: An international journal, 10(2), pp.201-202.

Bećirović, S., 2023. Fostering student engagement in implementing digital pedagogy. In Digital pedagogy: The use of digital technologies in contemporary education (pp. 97-111). Singapore: Springer Nature Singapore.

Belenkova, L.Y., Skudnyakova, Y.V. and Bosov, D.V., 2022. Digital pedagogy in the system of inclusive higher education. Interacción y Perspectiva: Revista de Trabajo Social, 12(1), pp.27-42.

Bhaskar, P. and Rana, S., 2024. The ChatGPT dilemma: unravelling teachers' perspectives on inhibiting and motivating factors for adoption of ChatGPT. Journal of information, Communication and Ethics in Society, 22(2), pp.219-239.

Bogner, A., Littig, B. and Menz, W., 2009. Introduction: Expert interviews—An introduction to a new methodological debate. In Interviewing experts (pp. 1-13). London: Palgrave Macmillan UK.

Bond, M., Bedenlier, S., Marín, V.I. and Händel, M., 2021. Emergency remote teaching in higher education: Mapping the first global online semester. International journal of educational technology in higher education, 18, pp.1-24.

Chakraborti, J., Dutta, A. and Jana, B., 2022. Digital marketing adoption among startups in India: A systematic review and bibliometric analysis. Journal of Information and Optimization Sciences, 43(6), pp.1301-1309.

Creely, E. and Laletas, S., 2020. Transitions, transformations and finding success. A phenomenological analysis of the experiences of a doctoral student in early candidature. Higher Education Research & Development, 39(3), pp.439-453.

Creswell, J.W. and Creswell, J.D., 2017. Research design: Qualitative, quantitative, and mixed methods approaches. Sage publications.

Croxall, B. and Koh, A., 2012. A Digital Pedagogy Unconference. Modern Language Association Convention. Boston. 3 January 2013 [online].
Davis, F.D., 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS quarterly, pp.319-340.
Dhakal, B.P., 2022. Digital pedagogy: An effective model for 21st century education. Academic Journal of Mathematics Education, 5(1), pp.1-9.
Eatough, V. and Smith, J.A., 2017. Interpretative phenomenological analysis. The Sage handbook of qualitative research in psychology, pp.193-209.
Efremova, N. and Huseynova, A., 2022, May. Digital pedagogy: Opportunities and challenges of learning in the information environment. In International Scientific Conference on Agricultural Machinery Industry "Interagromash"" (pp. 283-292). Cham: Springer International Publishing.
El-Hamamsy, L., Monnier, E.C., Avry, S., Chevalier, M., Bruno, B., Dehler Zufferey, J. and Mondada, F., 2024. Modelling the sustainability of a primary school digital education curricular reform and professional development program. Education and Information Technologies, 29(3), pp.2857-2904.
Engward, H. and Goldspink, S., 2020. Lodgers in the house: Living with the data in interpretive phenomenological analysis research. Reflective Practice, 21(1), pp.41-53.
Ertmer, P.A., Ottenbreit-Leftwich, A.T., Sadik, O., Sendurur, E. and Sendurur, P., 2012. Teacher beliefs and technology integration practices: A critical relationship. Computers & education, 59(2), pp.423-435.
Evans, D.N., 2023. A Qualitative Exploratory Case Study of Teachers' Perceptions Regarding the Implementation of Learner-Centered Pedagogy in Indian Elementary Schools (Doctoral dissertation, Northcentral University).
Fan, H., Li, B., Pasaribu, T. and Chowdhury, R., 2024. Online interviews as new methodological normalcy and a space of ethics: An autoethnographic investigation into Covid-19 educational research. Qualitative Inquiry, 30(3-4), pp.333-344.
Fatima, A., Verma, R., Singh, S., Kumar, M. and Srivastava, K., 2025. Digital Transformation and Continuance Intention of Online Learning Platforms Among Students in Higher Education Institutions in Remote India. In Perspectives on Digital Transformation in Contemporary Business (pp. 173-200). IGI Global Scientific Publishing.
Gerli, P., Clement, J., Esposito, G., Mora, L. and Crutzen, N., 2022. The hidden power of emotions: How psychological factors influence skill development in smart technology adoption. Technological Forecasting and Social Change, 180, p.121721.
Goarty, R. and Gupta, K.P., 2023. Digital transformation in Indian higher educational institutions: a qualitative exploration of administrators' perspectives. International Journal of Technology Enhanced Learning, 15(3), pp.286-310.
Goldspink, S. and Engward, H., 2019. Booming clangs and whispering ghosts: Attending to the reflexive echoes in IPA research. Qualitative Research in Psychology, 16(2), pp.291-304.
Government of India., 2020. National Education Policy 2020. Ministry of Education. https://www.education.gov.in/nep2020
Harper, C.V., McCormick, L.M. and Marron, L., 2024. Face-to-face vs. blended learning in higher education: a quantitative analysis of biological science student outcomes. International Journal of Educational Technology in Higher Education, 21(1), p.2.
Herbert, A., Nykvist, S., Hoa Man, B.T., Ta Thi, M.H., Do Thi Thuy, L. and Thi Anh Truong, T., 2022. Shifting identities and changing mindsets: A case of lecturers adopting digital pedagogies in Vietnam. Impacts of COVID-19 Pandemic's Distance Learning on Students and Teachers in Schools and in Higher Education: International Perspectives, pp.445-467.
Hernandez-Martinez, P. and Keane, T., 2025. Learning mathematics and its relevance through a digital storytelling assessment task at university. International Journal of Mathematical Education in Science and Technology, 56(5), pp.811-827.
Hu, S., Laxman, K. and Lee, K., 2020. Exploring factors affecting academics' adoption of emerging mobile technologies-an extended UTAUT perspective. Education and Information Technologies, 25(5), pp.4615-4635.
Huda, M., 2019. Empowering application strategy in the technology adoption: insights from professional and ethical engagement. Journal of Science and Technology Policy Management, 10(1), pp.172-192.
Imtiaz, M.A. and Mirhashemi, M.T., 2013, December. Analyzing trends in technology acceptance studies in education domain. In 2013 International Conference on Current Trends in Information Technology (CTIT) (pp. 23-27). IEEE.
Instefjord, E.J. and Munthe, E., 2017. Educating digitally competent teachers: A study of integration of professional digital competence in teacher education. Teaching and teacher education, 67, pp.37-45.
Kaushik, M.K. and Verma, D., 2020. Determinants of digital learning acceptance behavior: A systematic review of applied theories and implications for higher education. Journal of Applied Research in Higher Education, 12(4), pp.659-672.
Khan, I.A., 2021. Teaching of pharmacy in Saudi Arabia: Relevance of English and digital pedagogy. Bahrain Medical Bulletin, 43(3), pp.618-620.
Kim, J. and Lee, K.S.S., 2022. Conceptual model to predict Filipino teachers' adoption of ICT-based instruction in class: using the UTAUT model. Asia Pacific Journal of Education, 42(4), pp.699-713.
Kim, S., Wright, K., Westine, C.D., Lambert, M.C. and Sun, T., 2025. Examining Peer Observation Practices of Online Teaching and Learning Through Analysis of Existing Instruments. American Journal of Distance Education, 39(1), pp.3-27.

Kumalasari, R.A.D., Rahardjo, K., Kusumawati, A. and Sunarti, S., 2024. Biometric-based self-service technology adoption by older adult: empirical evidence from pension fund sector in Indonesia. Cogent Business & Management, 11(1), p.2325543.

Kumar, N., Singh, M., Upreti, K. and Mohan, D., 2021, June. Blockchain adoption intention in higher education: role of trust, perceived security and privacy in technology adoption model. In International Conference on Emerging Technologies and Intelligent Systems (pp. 303-313). Cham: Springer International Publishing.

Kumar, P., Kumar, P., Garg, R.K., Panwar, M. and Aggarwal, V., 2023. A study on teachers' perception towards E-learning adoption in higher educational institutions in India during the COVID-19 pandemic. Higher Education, Skills and Work-Based Learning, 13(4), pp.720-738.

Kumar, V. 2024. Digitalization Higher Education in India. Journal of Advanced Research in Education, 3(1), 14–17.

Kumar, V., 2019, May. The state of digital pedagogy in higher education. In Proceedings of International Conference on Digital Pedagogies (ICDP).

Kushagra, K. and Dhingra, S., 2021. An empirical analysis of the government cloud adoption in India. International Journal of Electronic Government Research (IJEGR), 17(3), pp.21-43.

Larkin, M., Shaw, R. and Flowers, P., 2019. Multiperspectival designs and processes in interpretative phenomenological analysis research. Qualitative research in psychology, 16(2), pp.182-198.

Laurillard, D., 2024. Trust the teachers: A collaborative approach to learning design solutions. In Envisioning the future of education through design (pp. 119-145). Singapore: Springer Nature Singapore.

Mannila, L., 2018, October. Digitally competent schools: teacher expectations when introducing digital competence in Finnish basic education. In Seminar. net (Vol. 14, No. 2, pp. 201-215).

Mateko, F.M., 2024. Opportunities in emerging technologies for Southern Africa: How the Global South should adopt to take advantage?. The Electronic Journal of Information Systems in Developing Countries, 90(5), p.e12321.

McCarthy, A., Maor, D. and McConney, A., 2017. Mobile technology in hospital schools: What are teachers' professional learning needs?. Journal of Technology and Teacher Education, 25(1), pp.61-89.

McCarthy, K.M., Glassburn, S.L. and Dennis, S.R., 2022. Transitioning to online teaching: A phenomenological analysis of social work educator perspectives. Social Work Education, 41(4), pp.641-659.

Milton, M. and Vozzo, L., 2013. Digital literacy and digital pedagogies for teaching literacy: Pre-service teachers' experience on teaching rounds. Journal of Literacy and Technology, 14(1), pp.72-97.

Ministry of Health and Family Welfare (MoHFW), Government of India (2021) *National Family Health Survey (NFHS-5), 2019–21: Phase II – India Fact Sheet*. New Delhi: MoHFW. Available at: NFHS-5 Phase II Report (Accessed: 21 March 2025).

Mishra, L., Gupta, T. and Shree, A., 2020. Online teaching-learning in higher education during lockdown period of COVID-19 pandemic. International journal of educational research open, 1, p.100012.

Montebello, M., 2017. Digital Pedagogies for Teachers' CPD. International Association for Development of the Information Society.

Morris, S.M., 2014. What is digital pedagogy? [blog] 10 March. Available at: https://www.seanmichaelmorris.com/what-is-digital-pedagogy

Moustakas, C., 1994. Phenomenological research methods. sage.

Nanjundaswamy, C., Baskaran, S. and Leela, M.H., 2021. Digital Pedagogy for Sustainable Learning. Shanlax International Journal of Education, 9(3), pp.179-185.

Negahban, A., 2024. Simulation in engineering education: The transition from physical experimentation to digital immersive simulated environments. Simulation, 100(7), pp.695-708.

Neyazi, T.A., 2020. Digital propaganda, political bots and polarized politics in India. Asian Journal of Communication, 30(1), pp.39-57.

Nizza, I.E., Farr, J. and Smith, J.A., 2021. Achieving excellence in interpretative phenomenological analysis (IPA): Four markers of high quality. Qualitative research in psychology, 18(3), pp.369-386.

OECD. (2021). *Education at a Glance 2021: OECD Indicators*. OECD Publishing. https://www.oecd.org/education/education-at-a-glance/

Osabutey, E.L., Senyo, P.K. and Bempong, B.F., 2024. Evaluating the potential impact of online assessment on students' academic performance. Information Technology & People, 37(1), pp.152-170.

Palinkas, L.A., Horwitz, S.M., Green, C.A., Wisdom, J.P., Duan, N. and Hoagwood, K., 2015. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. Administration and policy in mental health and mental health services research, 42(5), pp.533-544.

Patil, V.K., Patil, K.D., Gaikwad, S. and Yadav, S., 2023. QR code-based pedagogy for laboratory resource management in Indian higher education institutes. Journal of Engineering Education Transformations, pp.76-85.

Patton, M.Q., 1990. Qualitative evaluation and research methods. SAGE Publications, inc.

Peat, G., Rodriguez, A. and Smith, J., 2019. Interpretive phenomenological analysis applied to healthcare research. Evidence-based nursing, 22(1), pp.7-9.

Phutela, N. and Dwivedi, S., 2020. A qualitative study of students' perspective on e-learning adoption in India. Journal of applied research in higher education, 12(4), pp.545-559.

Pillai, R., Sivathanu, B., Metri, B. and Kaushik, N., 2024. Students' adoption of AI-based teacher-bots (T-bots) for learning in higher education. Information Technology & People, 37(1), pp.328-355.

Pittman, T. and Gaines, T., 2015. Technology integration in third, fourth and fifth grade classrooms in a Florida school district. Educational Technology Research and Development, 63(4), pp.539-554.
Pongsakdi, N., Kortelainen, A. and Veermans, M., 2021. The impact of digital pedagogy training on in-service teachers' attitudes towards digital technologies. Education and Information technologies, 26(5), pp.5041-5054.
Rahi, S., Othman Mansour, M.M., Alharafsheh, M. and Alghizzawi, M., 2021. The post-adoption behavior of internet banking users through the eyes of self-determination theory and expectation confirmation model. Journal of Enterprise Information Management, 34(6), pp.1874-1892.
Rahimi, I.D., Cohen Zilka, G., Finkelstein, I. and Cohen, R., 2024. The use of digital tools and digital learning skills in multicultural higher education campuses during the Covid-19 crisis. Israel Affairs, pp.1-18.
Rajasinghe, D., 2020. Interpretative phenomenological analysis (IPA) as a coaching research methodology. Coaching: An International Journal of Theory, Research and Practice, 13(2), pp.176-190.
Rana, A., Sarkar, B., Parida, R.K., Adhikari, S., Anandha Lakshmi, R., Akila, D. and Pal, S., 2023, September. A Data-Driven Analytical Approach on Digital Adoption and Digital Policy for Pharmaceutical Industry in India. In International Conference on Micro-Electronics and Telecommunication Engineering (pp. 509-521). Singapore: Springer Nature Singapore.
Rana, S., 2022. Exploring the student perspectives on the barriers of online learning during the COVID-19 pandemic: a qualitative study using interpretative phenomenological analysis. Journal of Applied Research in Higher Education, 14(4), pp.1554-1566.
Rapanta, C., Botturi, L., Goodyear, P., Guàrdia, L. and Koole, M., 2021. Balancing technology, pedagogy and the new normal: Post-pandemic challenges for higher education. Postdigital Science and Education, 3(3), pp.715-742.
Ray, A., Bala, P.K. and Dasgupta, S.A., 2019. Psychological analytics based technology adoption model for effective educational marketing. In Digital and Social Media Marketing: Emerging Applications and Theoretical Development (pp. 163-174). Cham: Springer International Publishing.
Rogers, E.M., 1962. Diffusion of innovations. New York: Free Press.
Rosamsi, S. and Nurdiani, N., 2024. Implementing the TPACK model to enhance learning outcomes of future biology teachers for the digital era. JPBIO (Jurnal Pendidikan Biologi), 9(2), pp.294-302.
Roy, D., 2022. Of pandemic and pedagogic reconstruction: an epistemological enquiry into digital education in India. In Evolution of Digitized Societies Through Advanced Technologies (pp. 45-61). Singapore: Springer Nature Singapore.
Sailin, S.N. and Mahmor, N.A., 2018. Improving student teachersâ€™ digital pedagogy through meaningful learning activities. Malaysian Journal of Learning and Instruction, 15(2), pp.143-173.
Savotina, N.A., Khachikyan, E.I., Bondarenko, V. and Ilyukhina, A., 2020, November. Digital technology in modern education: risks and resources. In Journal of Physics: Conference Series (Vol. 1691, No. 1, p. 012095). IOP Publishing.
Schmitz, M.L., Consoli, T., Antonietti, C., Cattaneo, A., Gonon, P. and Petko, D., 2024. Why do some teachers teach media literacy while others do not? Exploring predictors along the "will, skill, tool, pedagogy" model. Computers in Human Behavior, 151, p.108004.
Sharma, L. and Srivastava, M., 2020. Teachers' motivation to adopt technology in higher education. Journal of Applied Research in Higher Education, 12(4), pp.673-692.
Silvester, S., Sumarni, M.L. and Saputro, T.V.D., 2024. Pengaruh Kompetensi Technological Pedagogical Content Knowledge (TPACK) terhadap Keterampilan Guru dalam Mengimplemtasikan Pembelajaran Berbasis Digital. Journal of Education Research, 5(4), pp.4958-4965.
Singh, A., Sharma, S. and Paliwal, M., 2021. Adoption intention and effectiveness of digital collaboration platforms for online learning: the Indian students' perspective. Interactive Technology and Smart Education, 18(4), pp.493-514.
Singh, R. 2023. Digitization in Higher Education: Opportunities and Challenges. RESEARCH HUB International Multidisciplinary Research Journal, 10(4), 66–70. https://doi.org/10.53573/rhimrj.2023.v10n04.011
Smith, B. and McGannon, K.R., 2018. Developing rigor in qualitative research: Problems and opportunities within sport and exercise psychology. International review of sport and exercise psychology, 11(1), pp.101-121.
Smith, J.A. and Fieldsend, M., 2021. Interpretative phenomenological analysis. American Psychological Association.
Smith, J.A. and Osborn, M., 2007. Pain as an assault on the self: An interpretative phenomenological analysis of the psychological impact of chronic benign low back pain. Psychology and health, 22(5), pp.517-534.
Smith, J.A., Jarman, M. and Osborn, M., 1999. Doing interpretative phenomenological analysis. Qualitative health psychology: Theories and methods, 1(1), pp.218-240.
Smith, J.A., Larkin, M. and Flowers, P., 2021. Interpretative phenomenological analysis: Theory, method and research.
Souza, M.S. and Cardoso, V.C., 2024. Modelo TPACK, formação docente e Educação Matemática: relações, reflexões e possibilidades. Em Teia: Revista de Educação Matemática e Tecnológica Iberoamericana, 15(3), p.9.
Srideviponmalar, P., Wagle, S.A., Harikrishnan, R. and Sawant, R., 2022, October. Digital-Based Learning in Indian Government's Higher Education: Initiatives and Insights. In International Conference on Computing, Communications, and Cyber-Security (pp. 327-341). Singapore: Springer Nature Singapore.
Teo, T., Zhou, M. and Noyes, J., 2016. Teachers and technology: Development of an extended theory of planned behavior. Educational Technology Research and Development, 64(6), pp.1033-1052.
Tom Dieck, M.C., Cranmer, E., Prim, A. and Bamford, D., 2024. Can augmented reality (AR) applications enhance students' experiences? Gratifications, engagement and learning styles. Information Technology & People, 37(3), pp.1251-1278.

Tondeur, J., Scherer, R., Baran, E., Siddiq, F., Valtonen, T. and Sointu, E., 2019. Teacher educators as gatekeepers: Preparing the next generation of teachers for technology integration in education. British Journal of Educational Technology, 50(3), pp.1189-1209.

Tursunbayeva, A. and Gal, H.C.B., 2024. Adoption of artificial intelligence: A TOP framework-based checklist for digital leaders. Business Horizons, 67(4), pp.357-368.

Väätäjä, J.O. and Ruokamo, H., 2021. Conceptualizing dimensions and a model for digital pedagogy. Journal of Pacific Rim Psychology, 15, p.1834490921995395.

Van Schoors, R., Elen, J., Raes, A., Vanbecelaere, S. and Depaepe, F., 2023. The charm or chasm of digital personalized learning in education: Teachers' reported use, perceptions and expectations. TechTrends, 67(2), pp.315-330.

Vasileiou, K., Barnett, J., Thorpe, S. and Young, T., 2018. Characterising and justifying sample size sufficiency in interview-based studies: systematic analysis of qualitative health research over a 15-year period. BMC medical research methodology, 18(1), p.148.

Venkatesh, V., Morris, M.G., Davis, G.B. and Davis, F.D., 2003. User acceptance of information technology: Toward a unified view. MIS quarterly, pp.425-478.

Verdecchia, R., Lago, P. and De Vries, C., 2022. The future of sustainable digital infrastructures: A landscape of solutions, adoption factors, impediments, open problems, and scenarios. Sustainable Computing: Informatics and Systems, 35, p.100767.

Vicary, S., Young, A. and Hicks, S., 2017. A reflective journal as learning process and contribution to quality and validity in interpretative phenomenological analysis. Qualitative social work, 16(4), pp.550-565.

Wadmany, R. and Kliachko, S., 2014. The Significance of Digital Pedagogy: Teachers' Perceptions and the Factors Influencing Their Abilities as Digital Pedagogues. Journal of Educational Technology, 11(3), pp.22-33.

World Bank., 2022 a. *School enrollment, tertiary (% gross)*. World Development Indicators. https://data.worldbank.org

World Bank., 2022 b. *Individuals using the Internet (% of population)*. World Development Indicators. https://data.worldbank.org

Xu, K., Bao, X. and Lu, L., 2024. Elementary and secondary school students' perceptions toward the use of e-learning under the COVID-19 pandemic: a mixed-methods study. Information Technology & People, 37(1), pp.130-151.

Strategic Leverage Points in Blended Learning: A Systems Science Approach Using Grey-DEMATEL-ISM-MICMAC Framework in Higher Education

Xiaohan Liu and Pitipong Yodmongkol
College of Arts, Media and Technology, Chiang Mai University, Thailand

xiaohan_liu@cmu.ac.th
pitipong.y@cmu.ac.th (corresponding author)

https://doi.org/10.34190/ejel.24.2.4510

Abstract: E-learning has emerged as a cornerstone of contemporary higher education, offering flexible and technology-mediated environments that accommodate modern learning needs. Among its various modalities, blended learning (BL), which strategically integrates face-to-face and online instruction, has become a pivotal approach in higher education for enhancing learning outcomes and fostering talent cultivation. However, its successful implementation depends on the coordinated interaction of individual, technological, environmental, and course dimensions, constituting a complex network of interdependent factors that often remain fragmented in practice. Existing studies typically examine these factors in isolation and commonly rely on linear analytical approaches, providing limited insights into the systematic, comprehensive, and hierarchical understanding of the interrelationships among them. Understanding these structural interrelationships is therefore essential for identifying strategic leverage points that can optimise system performance and ensure the sustainable success of BL initiatives. To address this gap, this study proposes a systems science-based analytical framework that integrates the Grey Decision-Making Trial and Evaluation Laboratory (Grey-DEMATEL), Interpretive Structural Modelling (ISM), and Matrix Impact Cross Multiplication Applied to Classification (MICMAC). This integrated approach enables comprehensive and data-driven modelling of the causal parameters, hierarchical structure, and driving-dependence relationships among critical success factors of BL. First, ten critical success factors were identified through a systematic literature review and were then pairwise evaluated by twelve experts from a higher education institution in Thailand. Grey-DEMATEL was subsequently employed to quantify the causal properties and relative significance of these factors, while ISM was applied to construct a multi-layer hierarchical structure. MICMAC analysis further categorised the factors according to their driving and dependence powers. The results reveal a three-layer hierarchical structure of BL critical success factors, where policy support ($R - C = 2.78$), system quality ($R - C = 1.73$), and technical support ($R - \mathrm{C} = 1.62$) serve as key causal drivers, forming the institutional and technological foundation of the BL system. Course design and technology experience act as mediating linkages connecting institutional mechanisms with learning outcomes, while attitude, perceived usefulness, and interaction represent outcome-level indicators of system performance. Among these factors, course design exhibits the highest level of centrality value ($R + C = 18.6$) with the causal structure. The findings extend the understanding of the causal hierarchy and strategic leverage points for achieving BL success, illustrate how institutional and technological investment are realised through course design to improve individual experience. The study offers actionable insights for policymakers and instructional designers to inform data-driven decision-making and strategic planning in higher education, as well as how this is implemented at the level of the individual academic.

Keywords: Blended learning, Higher education, Multi-criteria decision analysis, Critical success factors, Systems science, Grey-DEMATEL-ISM-MICMAC

1. Introduction

The pervasive integration of digital technologies into daily life has not only profoundly reshaped how individuals live and interact but also transformed the creation and dissemination of knowledge (Liu, 2025). These ongoing transformations place increasing demands on higher education institutions (HEIs) to adapt proactively, not only in terms of delivery models, but also with respect to pedagogical approaches, instructional strategies, and teaching methodologies (Samala, Papadakis and Rawas, 2025). Within this context, blended learning has gained growing recognition for its potential to address diverse learning needs and enhance educational effectiveness (Mizza, Reese and Malouche, 2025), positioning it as a pivotal approach in shaping the educational frontier (Samala, Papadakis and Rawas, 2025; Verma et al., 2025). Blended learning (BL) is a pedagogical approach that integrates face-to-face instruction with online digital learning through appropriate alignment and balance (Castro, 2019). Stemming from the convergence of technological advancements and pedagogical theories, BL offers a flexible instructional paradigm that adapts to changing educational environments and accommodates diverse learning preferences. In recent years, a growing number of HEIs have embraced and implemented BL as

Cite this article: Liu, X. and Yodmongkol, P. 2026. "Strategic Leverage Points in Blended Learning: A Systems Science Approach Using Grey-DEMATEL-ISM-MICMAC Framework in Higher Education", *Electronic Journal of e-Learning*, 24(2), pp 112-131, https://doi.org/10.34190/ejel.24.2.4510

a key instructional strategy (Chen, 2022), recognising its capacity to foster lifelong learning and adaptability in an increasingly digitalised society (Dziuban et al., 2011).

As BL implementation matures, scholarly attention has shifted from adoption drivers to post-adoption effectiveness, particularly in identifying the factors that critically determine successful outcomes. Previous studies have explored various critical success factors, including course design quality (Huang, Kuang and Ling, 2022), classroom interaction (Majeed and Rehan Dar, 2022), the stability and information quality of learning management systems (Prifti, 2022), appropriate technological support (Su et al., 2023), and strong managerial support (Al-Mekhlafi et al., 2025; Zhao and Song, 2021). Empirical evidence suggests that the success of BL emerges from the interactions of multiple dimensions rather than from any single factor (Mielikäinen, 2022; Min and Yu, 2023). Nevertheless, most existing empirical studies rely primarily on linear analytical approaches such as structural equation modelling (SEM) and regression analysis (McCarthy and Palmer, 2023). While these methods are helpful for estimating associations among variables and revealing the significance of path effects, they are limited by linear assumptions and thus fail to capture potential nonlinear dynamics and feedback mechanisms among factors (Feng et al., 2024). Moreover, they typically examine single-layer relationships between independent and dependent variables (Hair et al., 2021), lacking the capacity to represent hierarchical structures or causal feedback loops. This methodological limitation prevents a thorough assessment of the status and structural interrelations of the factors within the system. In other words, the current research has yet to provide a systematic, comprehensive, and hierarchical understanding of the BL critical success factors in higher education. As technology-enhanced education continues to evolve, ongoing research remains necessary to deepen the understanding of effective teaching and learning practices and to support continuous refinement of BL initiatives (Samala, Papadakis and Rawas, 2025).

To address this gap, this study integrates Grey System theory with three complementary analytical methods, including Decision-Making Trial and Evaluation Laboratory (DEMATEL), Interpretive Structural Modelling (ISM), and Matrix-Based Cross-Impact Multiplication Applied to Classification (MICMAC), to construct a three-stage systemic framework that enables a comprehensive, hierarchical, and data-driven analysis of critical success factors in BL systems. By integrating these methods, the proposed framework provides a systematic approach to reveal the causal relationships, hierarchical structure, and driving-dependence mechanisms among critical success factors in BL systems. This integrated approach has been widely applied in complex system studies across various fields, including construction engineering (Zhang et al., 2024), supply chain management (Primadasa et al., 2025), and sustainable development (Bagherian et al., 2024), demonstrating its capacity to analyse systems characterised by numerous interacting factors, complex organisational interdependencies, and nonlinear causal feedback (Liu, Hu and Huang, 2024). In recent years, educational researchers have also increasingly employed such Multi-Criteria Decision Analysis (MCDA) techniques to explore interrelated factors in digital learning contexts, including virtual learning (Chuaphun and Samanchuen, 2024), simulation-based learning adoption (Asadi et al., 2024), and online learning quality (Zhou, Tang and Liu, 2025). However, prior studies have typically focused on simple causal analyses rather than a holistic exploration of system dynamics, and few studies have integrated Grey-DEMATEL, ISM, and MICMAC within a unified analytical framework to examine complex educational systems.

BL systems, by contrast, are inherently complex, multidimensional, and interdependent, requiring a systematic and comprehensive analytical approach to identify the cause-and-effect parameters and structural hierarchy relationships, and reveal the strategic leverage points necessary for effective improvement. To address this issue, the present study integrates three analytical methods into a unified framework that captures causal influence, hierarchical structure, and driving-dependence patterns. Accordingly, the following research questions are formulated:

RQ1: What causal relationships exist among the critical success factors of BL in higher education?

RQ2: What hierarchical structure characterises the relationships among these critical success factors?

RQ3: How can these critical success factors be categorised according to their driving-dependence powers, and which factors can be identified as strategic leverage points for improving BL effectiveness in higher education?

Specifically, Grey-DEMATEL enables the identification and evaluation of causal relationships and the strength of influence among factors, while accounting for uncertainty in expert judgments and providing a detailed visualisation of interdependencies (Bai and Sarkis, 2013). ISM utilises graph theory to partition complex systems into distinct elements (Feng et al., 2024) and constructs a multi-tiered structural model to enhance comprehension and analysis (Asadi et al., 2024). MICMAC analysis evaluates the driving and dependence powers of each factor and classifies them into four categories (Zhang et al., 2024), thereby identifying the most strategically valuable leverage points within the system (Almerino et al., 2024). This integrated approach enhances the analytical depth and interpretability of the results, providing a systemic and transparent understanding of the interconnections between critical success factors in BL environments.

This study extends existing research by employing Grey-DEMATEL, ISM, and MICMAC in an integrated manner within the educational context. The novelty of this work lies in its structured systems-modelling approach, which transcends linear analytical boundaries and reveals the complex causal chains and hierarchical structures underlying BL practices in HEIs. The results aim to provide a decision-support framework that is both systematic and actionable, enabling educational administrators and instructors to identify and prioritise strategic leverage points for improvement, thereby enhancing the overall effectiveness of BL initiatives.

2. Critical Success Factors of Blended Learning in Higher Education

The concept of critical success factors refers to the essential conditions or variables that determine the success of an organisation or system (Pollard and Cater-Steel, 2009). By taking critical success factors into account, organisations can identify the primary obstacles and prevent potential failures (Alkarney and Albraithen, 2018), and stakeholders can achieve better outcomes (Alqahtani and Rajkhan, 2020; Min and Yu, 2023). In essence, critical success factors are the elements that must be achieved for an organisation to be successful in attaining its desired goals (Selim, 2007). The critical success factors of BL are considered key influencing factors that significantly impact the effectiveness and outcomes of BL initiatives, which are reflected in the students' learning experience, performance, and satisfaction (Ghazal, Al-Samarraie and Aldowah, 2018; Min and Yu, 2023).

In order to determine the critical success factors of BL in HEIs, this study draws on the results of a systematic literature review (SLR) conducted by Liu and Yodmongkol (2023), following the approach outlined by Pattanasak et al. (2022). The relevant studies were retrieved from the Scopus database using the search string "blended-learning AND higher-education AND factor". The search was limited to conference proceedings and international journal publications between 2013 and 2023 to ensure research quality and relevance. Through the screening and refinement process, the initial set of 364 studies was narrowed down to 63 studies for in-depth analysis (Liu and Yodmongkol, 2023). The analysis focused on four dimensions based on the Complex Adaptive Blended Learning System (CABLS) framework (Wang, Han and Yang, 2015), including individual, technological, environmental, and course aspects. In addition to the SLR findings, recent international studies offer further context for understanding the critical success factors of BL. The evidence from higher education administrators and teachers across multiple European countries indicates that technological infrastructure, digital quality, and the availability of technical support and training are critical in enhancing BL effectiveness and student learning outcomes (Mohammadi, Paasivara and Kasurinen, 2025). At a broader level, bibliometric evidence from global higher education technology research suggests that technological and system quality, content quality, individual digital competence, and organizational-level technical support consistently underpin the successful implementation of technology-enhanced learning across diverse educational contexts (Samala, Papadakis and Rawas, 2025).

By synthesising the SLR findings with insights from recent international research, ten critical success factors with the highest frequency of occurrence within each dimension were identified, as presented in Table 1. Specifically, the individual dimension includes the competencies, perceptions, and attitudes of users involved in BL courses; the technological dimension encompasses the quality of digital systems and learning information; the environmental dimension reflects the institutional context, including policies and support mechanisms that facilitate BL implementation; and the course dimension represents the instructional design and interaction within BL courses. The selection of ten factors also aligns with the previous MCDA-based studies in educational contexts, which commonly analyse between 10 and 15 factors (Asadi et al., 2024; Chuaphun and Samanchuen, 2024) to maintain the analytical feasibility and interpretability of pairwise influence matrices.

Table 1: Critical success factors of blended learning in higher education

Factors	Code	Description	Reference
Computer self-efficacy	CSE	An individual's confidence in their capability to successfully utilize computers for educational tasks (Prifti, 2022).	(Katsarou, 2021; Prifti, 2022)
Technology experience	TE	The interaction and engagement individuals have with technology, involving the user's exposure to the system's functionality and the knowledge and skills the user gains from those interactions (Thompson, Compeau and Higgins, 2008).	(Al-Samarraie and Saeed, 2018; Alomari et al., 2020)

Factors	Code	Description	Reference
Perceived usefulness	PU	The degree to which the person perceives that blended learning, or its technology is a valuable and advantageous (Dakduk, Santalla-Banderali and Van Der Woude, 2018).	(Vo, Zhu and Diep, 2020; Wu and Liu, 2013)
Attitude	ATT	An individual's assessment of a specific behaviour, characterized by either favorable or unfavorable judgments (Wu et al., 2022).	(Acosta-Gonzaga and Ramirez-Arellano, 2021)
Interaction	INT	Timely and supportive communication that occurs during the learning process between teachers and students, and collaborative communication and activities that occur between peers in blended learning environment.	(Nortvig, Petersen and Balle, 2018; Taghizadeh and Hajhosseini, 2021)
System quality	SQ	Reliability, flexibility, integration, accessibility, timeliness, and integrity, which collectively ensure the system's effectiveness in meeting user needs and supporting their tasks (Li and Zhu, 2022).	(Majeed and Rehan Dar, 2022)
Information quality	INQ	The relevance, timeliness, accuracy, completeness, accessibility, adequacy, clarity, consistency, and format of content provided by an information system to its users (Ghazal, Al-Samarraie and Aldowah, 2018; Roca, Chiu and Martínez, 2006).	(Nikou and Maslov, 2023)
Policy support	PS	The establishment of guidelines, frameworks, and regulations that govern the implementation and operation of blended learning programs.	(Anthony et al., 2019; Galvis, 2018; Zhou, Smith and Al-Samarraie, 2023)
Technical support	TS	Support services for educators and learners in blended learning, encompassing training, guidance, and troubleshooting in technology use.	(Anthony et al., 2022; Bokolo Jr et al., 2020)
Course design	CD	The strategic creation and organization of learning content, technology, and activity to create quality learning environments and experiences for students.	(Huang, Kuang and Ling, 2022; Su et al., 2023)

3. Methodology

This study employed a structured system modelling approach that integrates Grey-DEMATEL, ISM, and MICMAC within the framework of systems science to analyse the BL critical success factors. This hybrid methodology enables a systematic examination of causal relationships and hierarchical structures among variables, thereby assisting decision-makers in identifying strategic leverage points for effective improvement. The methodological workflow is depicted in Figure 1.

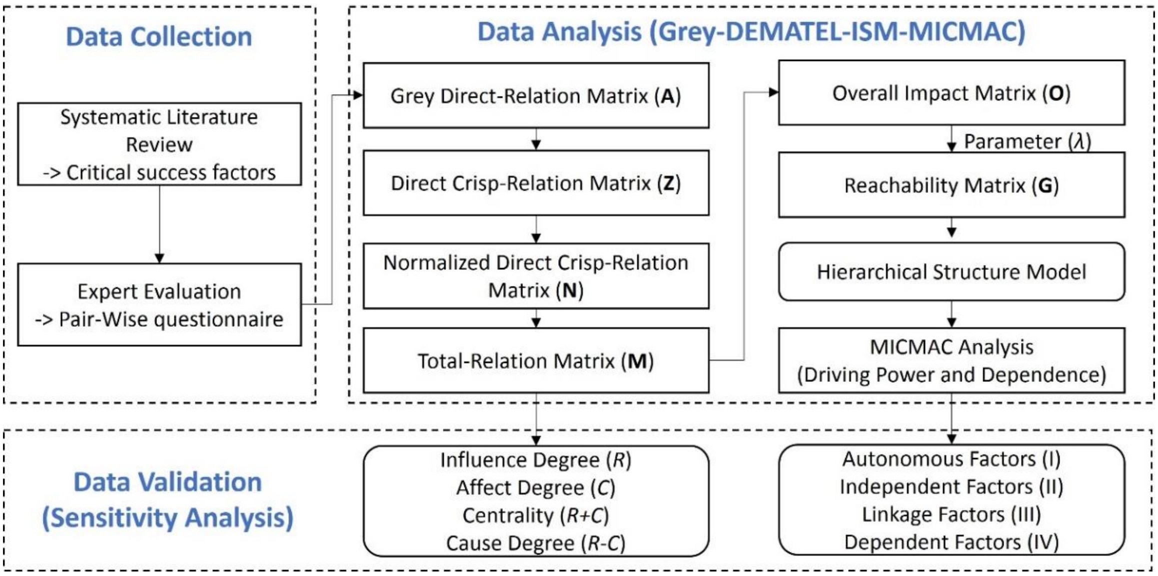


Figure 1: Workflow of Grey-DEMATEL-ISM-MICMAC methodology

The DEMATEL is a causal analysis technique based on expert judgment that identifies and visualises interrelationships among system variables (Aria, Jafari and Behifar, 2024). It quantifies the degree to which factors influence or are influenced by other factors through matrix computation and graphical representations (Dalvi-Esfahani et al., 2019), thereby capturing the causal relationships among system components (Aria, Jafari and Behifar, 2024) and the relatively significant levels of these relationships (Thakkar, 2021). It is considered a holistic and comprehensive method for analysing causal structures among factors (Lu, Huang and Wang, 2024). However, expert judgments are inevitably influenced by fuzziness and uncertainty in real-life situations (Liu et al., 2021), as well as by human bias, incomplete information, and inherent uncertainties (Pinili et al., 2024). To address these limitations, Grey system theory is incorporated into DEMATEL, which enhances the reliability, accuracy, and robustness of causal inference in uncertain environments (Deepu and Ravi, 2021). Despite its strength in revealing causal linkages, DEMATEL alone cannot intuitively depict the hierarchical structure of relationships within a system.

The ISM effectively complements this limitation by applying graph theory to decompose complex systems into structured, multi-level models (Feng et al., 2024). Through algebraic operations, ISM analyses the direct binary relationships among variables and applies Boolean logic to construct recursive and directed topologies (Lan et al., 2022). This process reveals the logical pathways linking foundational causal factors to higher-order outcome factors (Liu, Hu and Huang, 2024), thereby clarifying the hierarchical organisation and interdependence within the system. Nevertheless, the ISM is primarily oriented toward structural and macro-level exploration, and it does not quantify the relative contribution of each element to the system.

The MICMAC analysis further refines the ISM-derived structure by classifying factors according to their driving and dependence power (Du and He, 2025). Based on the principle of matrix multiplication, MICMAC categorises variables into four categories: autonomous, independent, linkage, and dependent, each exerting distinct effects on system dynamics, stability, and feedback (Bashir and Ojiako, 2020).

Taken together, the Grey-DEMATEL-ISM-MICMAC framework integrates the strengths of causal analysis, hierarchical modelling, and driving-dependence classification. This three-stage hybrid approach enables a systematic and holistic examination of the driving forces and interaction mechanisms within complex systems, thereby offering robust support for strategic decision-making.

3.1 Data Collection with Expert Evaluation

The first step involved identifying the factors or variables within the system as determined through the SLR and summarised in Table 1. An expert panel then conducted pairwise comparisons to assess the causal relationships among the ten critical success factors. In accordance with the data requirements of the DEMATEL method, an evaluation scale ranging from 0 to 4 was used to quantify causal influence (Pinili et al., 2024), and the corresponding grey values are presented in Table 2.

Table 2: Linguistic scale and corresponding grey values

Linguistic term	Influence score	Grey values
No influence	0	[0,0]
Very low influence	1	[0,0.25]
Low influence	2	[0.25,0.5]
High influence	3	[0.5,0.75]
Very high influence	4	[0.75,1]

The expert sampling technique was employed to select participants, ensuring that the data originated from individuals with specialised knowledge and relevant experience (Tuapawa, 2017). The specific criteria were established as follows: 1) individuals who hold a doctoral degree with more than five years of teaching experience in HEIs; 2) individuals who have conducted more than ten BL courses; and 3) individuals who have received teaching awards, honours, or other formal recognition for excellence in BL or e-learning. The invitation was issued to qualified experts via institutional email, accompanied by an official letter from the host organisation outlining the research objectives, procedures, and anticipated contributions. An information sheet and informed consent form were also provided, clearly stating the voluntary nature of participation, measures for confidentiality and data privacy, and the use of anonymised judgments for quantitative analysis and academic publication. Of the invited experts, twelve consented to participate, as detailed in Table 3. This sample size aligns with established methodological precedents in related studies, which have typically engaged between five and fifteen domain experts (Asadi et al., 2024; Khan et al., 2024).

Although all participating experts were affiliated with a single HEI in Thailand, prior MCDM studies indicate that the validity of expert judgments depends primarily on the diversity of experts' backgrounds, functional roles, experience, knowledge, and areas of specialisation (Du and Shen, 2024). This approach is consistent with Li and Xiao (2024) and Quiñones et al. (2020), who emphasise that multi-functional expert panels are more effective for identifying causal relationships in complex systems that help reduce bias and enhance the robustness of analytical results. As shown in Table 3, the expert panel in this study comprised members from different faculties and disciplinary backgrounds who occupied distinct functional roles, including instructional, management, and technical support roles. Several experts concurrently held multiple functional responsibilities, enabling cross-functional evaluation of causal relationships among critical success factors. In addition, this study incorporated grey theory and sensitivity analysis to further mitigate the uncertainty and subjectivity of expert judgments. As noted by Si et al. (2018), the integration of grey theory is particularly appropriate for systems that exhibit random uncertainty. Sensitivity analysis was subsequently conducted to assess the stability of the identified structure, thereby enhancing the robustness of the analytical results.

Table 3: Expert demographic information

Category	Sub-Category	Percentage
Position/ Role	Lecturer	8.33%
	Assistant Professor	50.00%
	Associate Professor	41.67%
Teaching Experience	5-15 years	33.33%
	16-25 years	16.67%
	More than 25 years	50.00%
Area of Expertise	E-learning/ Blended Learning	50.00%
	Educational Technology	25.00%
	Pedagogical Innovation	25.00%
Function Roles	Instructional	100.00%
	Management	50.00%
	Technical Support	25.00%

3.2 Data Analysis with Grey-DEMATEL-ISM-MICMAC Approach

The Grey-DEMATEL method was implemented in accordance with the methodology outlined by Raj and Sah (2019). The DEMATEL method analyses complex systems by identifying and evaluating pairwise relationships among a set of factors $x = \{x_i | i = 1,2, \dots ,10\}$. Grey systems are characterised by the use of grey numbers, grey equations, and grey matrices (Deepu and Ravi, 2021). In Grey system theory, $\otimes x$ represents a grey number, which is an interval defined by known lower upper $\overline{\otimes} x$ and lower $\underline{\otimes} x$ bound, while the distribution information of $\otimes x$ remains unknown. Specifically, $\otimes x$ is constrained within the range $[\underline{\otimes} x, \overline{\otimes} x]$, where $\underline{\otimes} x$ and $\overline{\otimes} x$ represent the lower and the upper limits, respectively. The influence scores of factors i on factors j $(\forall i, j)$ were obtained from the experts and then converted into corresponding grey values.

Step 1: the Grey direct-relation matrix (*A*) is calculated using equation (1), assigning equal weight to all twelve experts to ensure that each expert's judgment contributed equally to the aggregated evaluation.

$$\otimes x_{ij}^{12} = \left(\frac{\sum_{12} \underline{\otimes} x_{ij}^{12}}{12}, \frac{\sum_{12} \overline{\otimes} x_{ij}^{12}}{12} \right) \tag{1}$$

Step 2: the direct crisp-relation matrix (*Z*) is constructed using the Converting Fuzzy Data into Crisp Scores (CFCS) method (Opricovic and Tzeng, 2003), which transforms grey numbers into crisp values through a three-step procedure, as outlined below.

- Normalization

$$\Delta_{min}^{max} = \max_j \overline{\otimes} x_{ij} - \min_j \underline{\otimes} x_{ij} \tag{2}$$

$$\underline{\otimes}\,\overline{x}_{ij} = (\underline{\otimes}\,x_{ij} - \min_{j}\underline{\otimes}\,x_{ij})/\Delta_{min}^{max} \tag{3}$$

$$\overline{\otimes}\,\overline{x}_{ij} = (\overline{\otimes}\,x_{ij} - \min_{j}\underline{\otimes}\,x_{ij})/\Delta_{min}^{max} \tag{4}$$

- Calculating the normalized crisp value

$$y_{ij} = \frac{\underline{\otimes}\,\overline{x}_{ij}(1 - \underline{\otimes}\,\overline{x}_{ij}) + \overline{\otimes}\,\overline{x}_{ij} \cdot \overline{\otimes}\,\overline{x}_{ij}}{1 - \underline{\otimes}\,\overline{x}_{ij} + \overline{\otimes}\,\overline{x}_{ij}} \tag{5}$$

- Computing the final crisp value

$$z_{ij} = \min_{j}\underline{\otimes}\,x_{ij} + y_{ij}\,\Delta_{min}^{max} \tag{6}$$

Step 3: the normalized direct crisp-relation matrix (*N*) is computed by equation (7).

$$N = \frac{1}{max\left(\max\limits_{1\le i\le n}\sum_{j=1}^{n}x_{ij}\,,\max\limits_{1\le j\le n}\sum_{i=1}^{n}x_{ij}\right)} \times Z \tag{7}$$

Step 4: the total-relation matrix (*M*) is constructed and calculated using equation (8), where I is the identity matrix.

$$M = N(I - N)^{-1} \tag{8}$$

Step 5: categorise the elements into net cause and net effect groups using (*R* – *C*), compute *R* and *C* using formulas (9) and (10), separately. The (*R* + *C*) vector, referred to as the "centrality" vector, indicates the relative importance of all elements. The elements in the (*R* – *C*) vector, known as the "cause-degree" vector, are categorised into the net cause group if $R_i - C_j > 0$, while those with $R_i - C_j < 0$ are placed in the net effect group.

$$R = \sum\nolimits_{j=1}^{n} M_{ij} \forall i \tag{9}$$

$$C = \sum\nolimits_{i=1}^{n} M_{ij} \forall j \tag{10}$$

Step 6: the ISM-MICMAC was implemented as outlined by Zhang et al. (2024). The parameter (λ) is calculated using equation (11) to filter out insignificant relationships and simplify the system structure, which is the addition formula of the mean values of M_{ij} and the standard deviation to the mean. This threshold setting is commonly adopted in ISM studies to improve the accuracy of the results of calculating the reachability matrix (Hu et al., 2024).

$$\lambda = \mu + \sigma \tag{11}$$

Step 7: the overall impact matrix (*O*) is constructed using equation (12), where I is the identity matrix.

$$O = M + I \tag{12}$$

Step 8: the overall impact matrix (*O*) is transformed into the reachability matrix (*G*) using equation (13).

$$G_{ij} = \begin{cases} 1, O_{ij} \ge \lambda \ (i,j = 1,2,\ldots,10) \\ 0, O_{ij} \ge \lambda \ (i,j = 1,2,\ldots,10) \end{cases} \tag{13}$$

Step 9: the multi-level hierarchical structure model is constructed using equation (14).

$$S_{(x_i)} = S_{(x_i)} \cap Q_{(x_i)} \tag{14}$$

It involves systematically decomposing the interrelationships among factors and revealing the internal hierarchy within the system. This procedure follows the principles of interval and inter-level decomposition, where system elements are first divided into distinct subsets and subsequently organised into hierarchical levels according to their relational dependencies. The process begins with the reachability matrix (*G*), which represents the direct and indirect relationships among factors derived from the previous analytical stage. For each factor x_i, two sets are identified: the reachability set $S_{(x_i)}$, which consists of all factors x_j for which $O_{ij} = 1$, representing those that can be reached from x_i; and the antecedent set $Q_{(x_i)}$, which consists of all factors x_j for which $O_{ji} = 1$, representing those that can reach x_i. The intersection of these two sets, $S_{(x_i)} \cap Q_{(x_i)}$, determines the hierarchical position of x_i. When the reachability set $S_{(x_i)}$ and its intersection are identical, that is $S_{(x_i)} = S_{(x_i)} \cap Q_{(x_i)}$, the factor x_i is assigned to the highest hierarchical level, as it no longer influences any other unassigned elements in the system. Once the factors at the highest level are identified, their corresponding rows and columns are removed from the reachability matrix (*G*). The same procedure is then repeated iteratively for the remaining factors until all elements are allocated to specific hierarchical levels.

Step 10: the driving power ($D_{(i)}$) and dependency value ($P_{(i)}$) of each factor are calculated using formulas (15) and (16), which is the sum the rows and columns of the reachability matrix (*O*), respectively. And then use the average value of dependence and driving power values as the dividing line of the quadrant and divide the factors into four quadrants: Autonomous (I), Independent (II), Linkage (III), and Dependent (IV).

$$D_{(i)} = \sum_{j=1}^{n} G_{ij} \tag{15}$$

$$P_{(i)} = \sum_{i=1}^{n} G_{ji} \tag{16}$$

3.3 Data Validation with Sensitivity Analysis

The results obtained from the data analysis with Grey-DEMATEL-ISM-MICMAC may be influenced by biases due to varying expertise and experience among the experts (Raj and Sah, 2019). To mitigate such effects and enhance the reliability and accuracy of the findings, sensitivity analysis was employed. Sensitivity analysis is a technique commonly used to examine how changes in a model's inputs influence the uncertainty in its outputs (Saltelli et al., 2008). This analysis involves adjusting the weight of the particular experts while maintaining uniform weights for the others to assess the overall effect on the system (Xia, Govindan and Zhu, 2015). This analytical framework finds extensive application in MCDA to ensure the robustness and reliability of the results (Więckowski and Sałabun, 2023).

The effects of these variations on the causal structure (Grey-DEMATEL), hierarchical levels (ISM), and driving-dependence classifications (MICMAC) were examined. By verifying that no single expert exerts disproportionate influence on the result, this validation step ensures the robustness of the analysis and strengthens the credibility of the findings.

4. Data Analysis and Result

4.1 Grey-DEMATEL Result

The Grey-DEMATEL technique was first applied to determine the causal relationships and influence strength among the identified critical success factors. The experts' judgments were aggregated to construct the initial Grey direct-relation matrix (*A*) using Grey system operations, as presented in Table 4, which was subsequently converted into crisp value through the CFCS method. The influence degree (*R*) and affect degree (*C*) were computed according to equations (3)-(10) to quantify the impact of each factor on others and vice versa.

Table 4: Grey direct-relation matrix (*A*)

	CSE	TE	PU	ATT	INT	SQ	INQ	PS	TS
CSE	[0.00, 0.00]	[0.60, 0.85]	[0.63, 0.88]	[0.52, 0.77]	[0.48, 0.73]	[0.31, 0.54]	[0.35, 0.60]	[0.29, 0.52]	[0.40, 0.63]
TE	[0.69, 0.94]	[0.00, 0.00]	[0.60, 0.85]	[0.58, 0.83]	[0.60, 0.85]	[0.40, 0.60]	[0.44, 0.65]	[0.29, 0.50]	[0.46, 0.67]
PU	[0.50, 0.75]	[0.56, 0.81]	[0.00, 0.00]	[0.65, 0.90]	[0.52, 0.77]	[0.40, 0.60]	[0.42, 0.65]	[0.38, 0.58]	[0.42, 0.63]
ATT	[0.54, 0.79]	[0.54, 0.79]	[0.60, 0.85]	[0.00, 0.00]	[0.58, 0.83]	[0.33, 0.54]	[0.35, 0.56]	[0.29, 0.52]	[0.31, 0.52]

	CSE	TE	PU	ATT	INT	SQ	INQ	PS	TS
INT	[0.46, 0.71]	[0.46, 0.71]	[0.50, 0.75]	[0.54, 0.79]	[0.00, 0.00]	[0.40, 0.60]	[0.52, 0.75]	[0.29, 0.50]	[0.40, 0.60]
SQ	[0.48, 0.71]	[0.58, 0.83]	[0.56, 0.81]	[0.56, 0.81]	[0.58, 0.83]	[0.00, 0.00]	[0.56, 0.79]	[0.44, 0.69]	[0.54, 0.79]
INQ	[0.40, 0.63]	[0.42, 0.64]	[0.56, 0.81]	[0.56, 0.81]	[0.48, 0.73]	[0.35, 0.58]	[0.00, 0.00]	[0.27, 0.50]	[0.29, 0.52]
PS	[0.44, 0.69]	[0.50, 0.75]	[0.54, 0.79]	[0.50, 0.75]	[0.46, 0.71]	[0.65, 0.90]	[0.48, 0.71]	[0.00, 0.00]	[0.69, 0.94]
TS	[0.58, 0.83]	[0.63, 0.88]	[0.56, 0.81]	[0.54, 0.79]	[0.52, 0.77]	[0.54, 0.77]	[0.48, 0.73]	[0.44, 0.69]	[0.00, 0.00]
CD	[0.54, 0.79]	[0.56, 0.81]	[0.65, 0.90]	[0.63, 0.88]	[0.67, 0.92]	[0.52, 0.75]	[0.54, 0.77]	[0.46, 0.71]	[0.54, 0.77]

As shown in Table 5, the factors with the highest influence degrees (*R*) include course design (CD), technical support (TS), and system quality (SQ), indicating that these elements exert substantial influence on other factors within the BL system. Conversely, the factors with the highest dependence degrees (*C*) include course design (CD), attitude (ATT), and interaction (INT), suggesting that these factors are most affected by others. Notably, CD exhibited both the highest influence and dependence values, highlighting its dual role as a central driver and recipient of systemic interactions.

Table 5: Grey-DEMATEL analysis result

	Influence Degree (*R*)	Affect Degree (*C*)	Centrality (*R* + *C*)	Cause-Degree (*R* - *C*)	Rank	Properties
CSE	7.67	8.42	16.09	-0.75	6	Effect
TE	8.43	8.74	17.16	-0.31	3	Effect
PU	8.08	9.31	17.39	-1.23	2	Effect
ATT	7.60	9.14	16.74	-1.54	4	Effect
INT	7.55	8.87	16.42	-1.33	5	Effect
SQ	8.73	7.00	15.73	1.73	8	Cause
INQ	7.01	7.46	14.47	-0.45	10	Effect
PS	8.70	5.92	14.62	2.78	9	Cause
TS	8.82	7.20	16.02	1.62	7	Cause
CD	9.04	9.57	18.60	-0.53	1	Effect

The causal degree (*R* – *C*) was further used to classify factor properties, identifying three causal factors and seven effect factors. The cause-effect relationships were visualised in Figure 2 to further clarify the connections. The factors above the horizontal axis, with a positive causal degree, represent causal drivers that exert a substantial influence on the system. In contrast, the factors below the axis, with a negative causal degree, are primarily affected by others. The centrality value (*R* + *C*) reflects the relative prominence of each factor, with higher scores indicating stronger systemic significance. The top three factors were course design (CD), perceived usefulness (PU), and technology experience (TE).

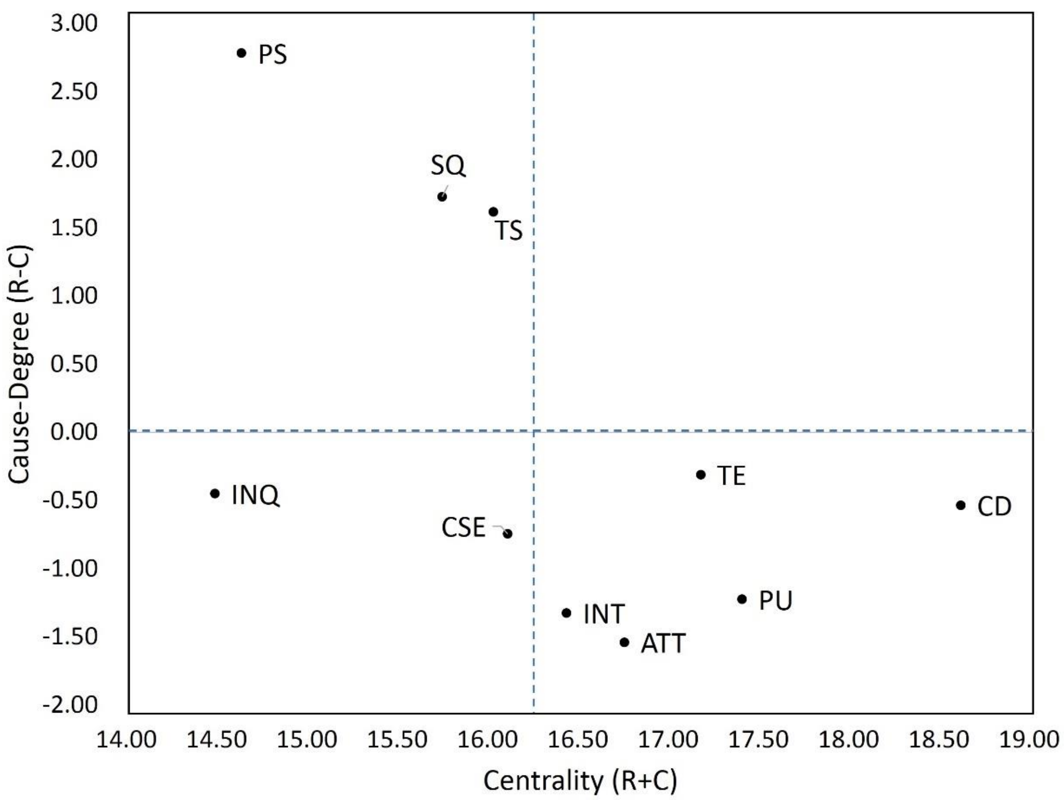


Figure 2: The cause-and-effect diagram

4.2 ISM Result

Based on the total-relation matrix (*M*) derived from Grey-DEMATEL, the ISM procedure was applied to construct a hierarchical structure of critical success factors. The reachability matrix (*G*) was constructed using equation (13) with the parameter (λ = 0.95) as outlined in Table 6.

Table 6: Reachability matrix (*G*)

	CSE	TE	PU	ATT	INT	SQ	INQ	PS	TS	CD
CSE	1	0	0	0	0	0	0	0	0	0
TE	0	1	1	1	0	0	0	0	0	1
PU	0	0	1	0	0	0	0	0	0	1
ATT	0	0	0	1	0	0	0	0	0	0
INT	0	0	0	0	1	0	0	0	0	0
SQ	0	0	1	1	1	1	0	0	0	1
INQ	0	0	0	0	0	0	1	0	0	0
PS	0	0	1	1	0	0	0	1	0	1
TS	0	1	1	1	1	0	0	0	1	1
CD	0	1	1	1	1	0	0	0	0	1

The reachability set ($S_{(x_i)}$) and antecedent set ($Q_{(x_i)}$) were derived for each factor through interval and inter-level decomposition, and intersections were used to determine factor levels iteratively. The factors whose reachability-antecedent intersection contained only themselves were classified at the top level, while remaining factors were processed as outlined in equation (14) until all were assigned to levels, as summarised in Table 7.

Table 7: The hierarchical set analysis

Factors	Reachability Set ($S_{(x_i)}$)	Antecedent Set ($Q_{(x_i)}$)	Intersection	Hierarchy
CSE	CSE	CSE	CSE	Top Layer
TE	TE, PU, ATT, CD	TE, TS, CD	TE, CD	Middle Layer
PU	PU, CD	TE, PU, SQ, PS, TS, CD	PU, CD	Top Layer
ATT	ATT	TE, ATT, SQ, PS, TS, CD	ATT	Top Layer
INT	INT	INT, SQ, TS, CD	INT	Top Layer
SQ	PU, ATT, INT, SQ, CD	SQ	SQ	Bottom Layer
INQ	INQ	INQ	INQ	Top Layer
PS	PU, ATT, PS, CD	PS	PS	Bottom Layer
TS	TE, PU, ATT, INT, TS, CD	TS	TS	Bottom Layer
CD	TE, PU, ATT, INT, CD	TE, PU, SQ, PS, TS, CD	TE, PU, CD	Middle Layer

The resulting multi-level hierarchical model organised the ten critical success factors into three distinct layers, as presented in Figure 3, which clearly illustrates the flow from root causes to intermediate factors and outcomes. The bottom layer represents the strategic level, comprising elements that exert long-term and structural influence on upper-level factors. These include system quality (SQ), policy support (PS), and technical support (TS). The middle layer bridges the strategic and outcome-oriented levels, including technology experience (TE) and course design (CD). The top layer encompasses the outcome-level factors, including computer self-efficacy (CSE), perceived usefulness (PU), attitude (ATT), interaction (INT), and information quality (INQ). This hierarchical structure provides clear insight into which factors serve as foundational strategic drivers and which are dependent outcomes, facilitating strategic prioritisation for BL initiative improvement.

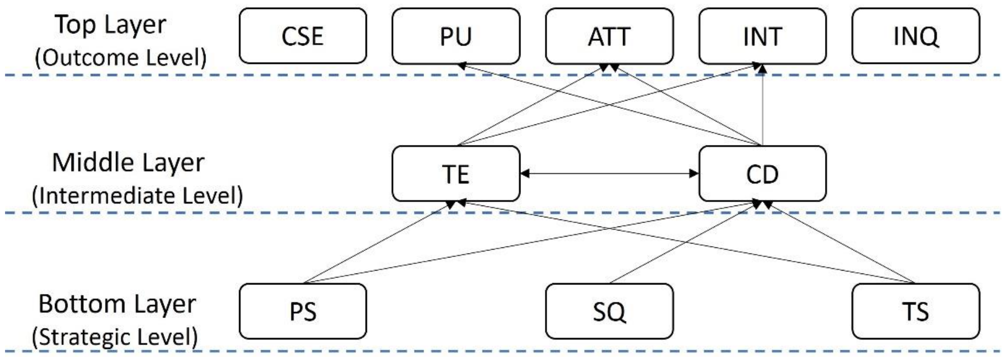


Figure 3: The multi-level structure diagram

4.3 MICMAC Result

MICMAC analysis was conducted to classify the BL critical success factors according to their driving power ($D_{(i)}$) and dependency value ($P_{(i)}$). Using the reachability matrix (*G*) from ISM, the $D_{(i)}$ and $P_{(i)}$ were calculated using equations (15) and (16), respectively, and the factors were subsequently categorised into four groups, as illustrated in Figure 4.

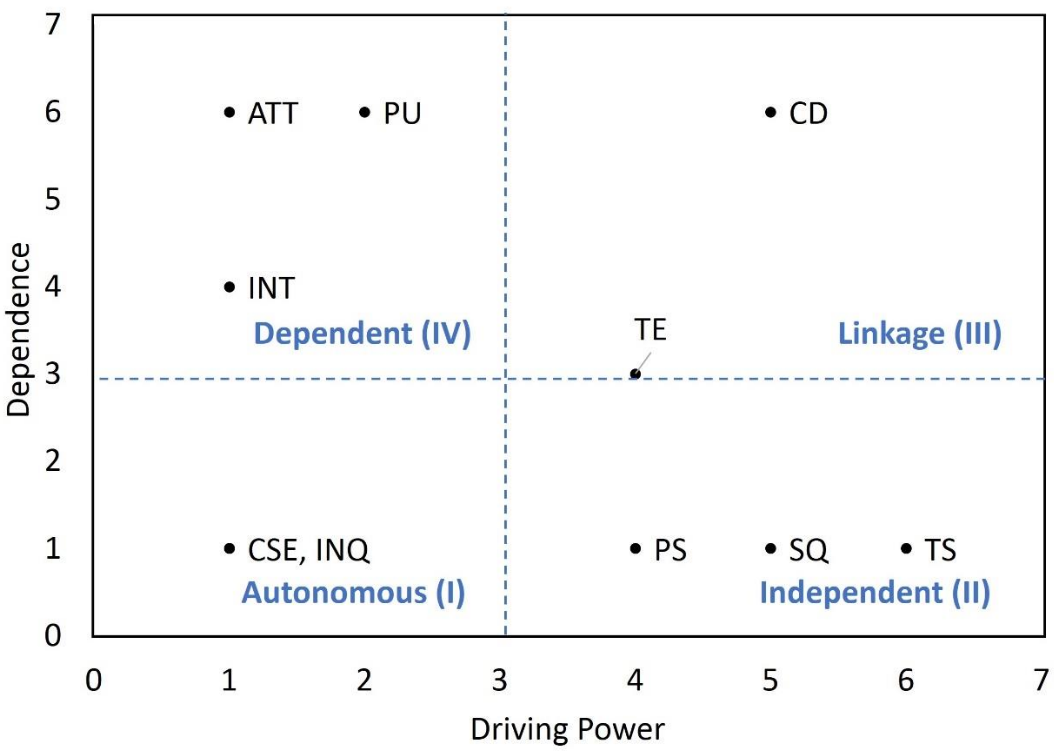


Figure 4: The multi-level structure diagram

The MICMAC analysis provides further insights into the driving and dependence relationships among critical success factors. Autonomous factors in quadrant I exhibit low driving and low dependence power, meaning they have little influence on other factors and are not significantly influenced by them. This group includes computer self-efficacy (CSE) and information quality (INQ), which tend to exert indirect or marginal effects on the outcome. Independent factors in quadrant II are characterised by strong driving power and weak dependence, exerting a strong influence on other factors in the system but are not heavily influenced by them. This quadrant includes policy support (PS), system quality (SQ), and technical support (TS), which are the factors that provide essential structural and operational support for BL practice. Linkage factors in quadrant III possess both high driving and dependence power, indicating that they both influence and are influenced by other factors in the system. They often act as crucial connectors between elements. This category comprises technology experience (TE) and course design (CD). Dependent factors in quadrant IV exhibit strong dependence but weak driving power. They are typically located in the upper layers of the ISM hierarchy, indicating that other factors greatly influence them but do not have a significant impact on the system. These include attitude (ATT), perceived usefulness (PU), and interaction (INT).

4.4 Sensitivity Analysis

Sensitivity analysis was performed to assess the robustness of the integrated Grey-DEMATEL-ISM-MICMAC results. The analysis compared the outcomes under different expert weighting schemes, where experts were categorised into four groups according to their teaching experience and domain expertise. By systematically varying the relative weights assigned to these groups, four experimental scenarios were developed, as summarised in Table 8.

Table 8: Different group weights assignment for sensitivity analysis

	Group 1	Group 2	Group 3	Group 4
Scenario 1	0.34	0.22	0.22	0.22
Scenario 2	0.22	0.34	0.22	0.22
Scenario 3	0.22	0.22	0.34	0.22
Scenario 4	0.22	0.22	0.22	0.34

For each scenario, the weighted aggregation of expert judgments was recalculated, and the corresponding causal relationships, hierarchical levels, and driving-dependence classifications were reanalysed, as presented in Table 9. The results demonstrated a high level of consistency across all scenarios, with no substantial differences observed in the causal structure, hierarchical layers, or factor classifications. This consistency suggests that experts with different levels of teaching experience and domain expertise share broadly similar perceptions towards the relationships of BL critical success factors. These findings confirm that the model's outcomes are structurally robust and are not significantly affected by variations in expert weighting schemes, thereby enhancing the credibility of the analytical results.

Table 9: Sensitivity analysis result

	DEMATEL (Cause-Effect)				ISM (Hierarchy Level)				MICMAC (Classification)			
	S1	S2	S3	S4	S1	S2	S3	S4	S1	S2	S3	S4
CSE	-0.67	-0.82	-0.79	-0.82	Top	Top	Top	Top	I	I	I	I
TE	-0.36	-0.20	-0.35	-0.34	Middle	Middle	Middle	Middle	III	III	III	III
PU	-1.29	-1.13	-1.22	-1.19	Top	Top	Top	Top	IV	IV	IV	IV
ATT	-1.55	-1.44	-1.63	-1.48	Top	Top	Top	Top	IV	IV	IV	IV
INT	-1.43	-1.17	-1.61	-1.12	Top	Top	Top	Top	IV	IV	IV	IV
SQ	1.93	1.58	1.85	1.60	Bottom	Bottom	Bottom	Bottom	II	II	II	II
INQ	-0.68	-0.56	-0.43	-0.30	Top	Top	Top	Top	I	I	I	I
PS	2.84	2.72	2.91	2.63	Bottom	Bottom	Bottom	Bottom	II	II	II	II
TS	1.83	1.58	1.85	1.52	Bottom	Bottom	Bottom	Bottom	II	II	II	II
CD	-0.64	-0.55	-0.57	-0.50	Middle	Middle	Middle	Middle	III	III	III	III

5. Discussion

This study systematically analysed the cause-effect parameter, hierarchical structure, and driving-dependence mechanisms among BL critical success factors in HEIs using the integrated Grey-DEMATEL-ISM-MICMAC approach.

Cause-effect parameter: The Grey-DEMATEL results identified policy support (PS), system quality (SQ), and technical support (TS) as primary causal drivers that establish the institutional and technological conditions necessary for BL success. Among these, PS was identified as the most influential factor, offering institutional strategic direction for resource allocation and technological integration (Chen et al., 2024). The effective PS ensures the infrastructure development and platform stability, promotes user-centred design, and fosters collaboration among stakeholders (Laohajaratsang, 2017). SQ constitutes the second major causal driver by maintaining stable performance, intuitive interfaces, and convenient features of learning platforms (Katsarou, 2021), while efficient TS assists users in resolving technical difficulties, thereby minimising participation barriers and sustaining online learning activities (Feng, He and Ding, 2023). These success factors collectively establish a supportive environment that enables pedagogical and technological advancements. Notably, course design (CD) exhibited the highest level of centrality in the causal structure, indicating its strong interconnectedness with other factors in the system (Wang et al., 2025). This implies that HEIs should place greater emphasis on ensuring coherent alignment among course content, instructional strategies, and technological features to strengthen the interaction and enhance cognitive engagement (Law, Geng and Li, 2019).

Hierarchical structure: The ISM analysis revealed a distinct three-layer hierarchy that aligns closely with the causal findings. The foundational layer comprises policy support (PS), system quality (SQ), and technical support (TS), forming the structural base of the BL system. Technology experience (TE) and course design (CD) constitute the intermediate layer, representing the mechanisms that transform institutional and technological resources into pedagogical processes. This finding is consistent with the results of the DEMATEL analysis, where CD exhibited the highest values of both influence and dependence. A well-designed BL course not only shapes learners' perceptions and experiences at the outcome level but is also influenced by enabling factors at the strategic level. This suggests that prioritising course design capacity within BL development strategies is essential to ensure that institutional and technological resources are effectively integrated into pedagogical practice, ultimately leading to improved learning performance. The top layer comprises outcome-oriented elements, including perceived usefulness (PU), attitude (ATT), and interaction (INT), which reflect the learners' behavioural and cognitive responses. This multi-layered structure illustrates a logical progression from institutional and technological support to individual-level engagement and perception. This finding is consistent with prior studies (Arjanto and Telussa, 2024), which indicate that socio-economic support and the provision of adequate technological infrastructure enhance the quality of pedagogical design, thereby improving student participation, engagement, and learning outcomes. Interestingly, computer self-efficacy (CSE) and information quality (INQ) appear relatively independent within the hierarchical model, revealing that they have a limited influence within the current system structure. In other words, these two variables exhibit relatively weak causal connections with other factors in the BL system: they are neither strongly influenced by foundational institutional support nor directly associated with students' behavioural and cognitive responses at the outcome level. This suggests that their influence may operate more indirectly through pedagogical mechanisms or other intermediary elements. Moreover, the bidirectional relationship between CD and TE suggests the presence of a self-reinforcing mechanism within the BL process. In other words, these factors mutually shape and reinforce one another. This observation aligns with the findings of Konstantinidou and Nisiforou (2022), who argued that well-structured and logically sequenced course design promotes higher technological engagement and improved learning outcomes. These technological improvements subsequently feed back into the instructional process, contributing to higher-quality design practices.

Driving-dependence classification: The MICMAC analysis further substantiates the structural interpretation by categorising factors according to their driving and dependence power. Through this classification, the analysis clarifies their functional roles within the causal network and helps to establish strategic intervention priorities (Javan Jafari Bojnordi et al., 2025). PS, SQ, and TS are categorised as independent factors, serving as fundamental strategic levers for achieving BL success. This result is consistent with the patterns identified in the Grey-DEMATEL and ISM analyses, underscoring the necessity for HEIs to strategically invest in these high-driving elements to ensure the sustainable advancement of BL initiatives. Without adequate investment in those enabling conditions, technology-enhanced learning environments are unlikely to achieve sustained effectiveness. TE and CD act as linkage factors, connecting institutional inputs and learning outcomes. This implies that they serve as key transformative mechanisms through which institutional investments are converted into meaningful learning experiences. This helps explain why many HEIs experience limited success in implementing BL initiatives (Sareen and Mandal, 2024): even when supportive policies, technical support mechanisms, and stable systems are well-provided, BL practices are often implemented as a simple juxtaposition of face-to-face instruction and online components (Rasheed, Kamsin and Abdullah, 2020; Rix, 2011), leading to constrained learning experiences. This finding indicates that the instructional effectiveness of BL courses depends on instructors' capacity to reconfigure the online and offline instructional structures in line with system capabilities and level of technical readiness (Ren et al., 2025). Consequently, providing targeted professional development for instructors in BL design and technological competencies plays a substantial role in supporting instructional effectiveness, and ultimately, improving students' learning experiences. For the administrators and policymakers of HEIs, fostering cross-functional collaboration among educators, instructional designers, and technologists is essential for aligning pedagogical objectives with technological capabilities to maximise the success of the BL initiatives. Conversely, ATT, PU, and INT are identified as dependent factors, reflecting how upstream elements collectively shape user experiences and learning results. These variables serve as evaluation indicators for monitoring the effectiveness of BL, enabling HEIs to iteratively refine their policies and strategic orientations through continuous assessment. Overall, this classification clarifies a coherent strategic framework that independent factors constitute the principal targets for institutional investment, linkage factors function as systemic connectors that transform the organisational investments into learning effectiveness, dependent factors act as outcome-based monitoring indicators, and autonomous factors occupy marginal positions within the system. These findings complement previous BL research (Hua, Wang and Li, 2024; Jannat Nipa and Hoque, 2025) that has applied linear analytical approaches to examine predictive relationships among variables, primarily focusing on what factors matter and the strength of their effects on learning outcomes. In contrast, this study moves beyond a linear causal perspective by demonstrating how these factors interact within the system to influence BL performance. This structured perspective enables HEIs to avoid fragmented improvement strategies and supports the BL progression toward more systematic and sustainable development.

By synthesising the three analytical results, this study conceptualises BL as a dynamic system shaped by multi-level interactions among institutional policies, technological affordances, and pedagogical practices. The findings provide practical implications and actionable insights for higher education administrators and policymakers seeking to enhance the effective management and sustainability of BL initiatives. By leveraging the identified causal and hierarchical structures, decision-makers can more effectively prioritise key areas for improvement. In particular, the institutional and technological dimensions should be recognized as the structural basis of the

BL system and serve as the foundational starting points for systemic optimisation; whereas course design and technology experience act as key leverage points for continuous enhancement. These mediating mechanisms transform institutional and technological drivers into pedagogical effectiveness. These findings also suggest that achieving efficiency, resilience, and long-term sustainability of BL systems in post-pandemic e-learning development requires strategic investment in these high-driving-power factors, as well as a systematic approach that aligns institutional governance, digital infrastructure, and pedagogical practice. In practical terms, HEIs should establish clear institutional guidelines and governance frameworks for the implementation of BL initiatives, ensuring that institutional support mechanisms, digital infrastructure, and technical resources are systematically aligned with course development needs. Institutions should also communicate these policies effectively to faculty and provide sustained technical and instructional support to facilitate the design and delivery of high-quality BL courses. Furthermore, improving system quality and ensuring responsive technical support are essential for maintaining reliable and user-friendly learning environments that empower instructors to design interactive and engaging learning experiences

This study contributes to the literature by introducing a systems science-based analytical framework into educational research, integrating causal, hierarchical, and driving-dependence analyses. This systemic perspective moves beyond factor-based explanations toward a structural understanding of BL success. The integrative approach advances the understanding of how individual, technological, environmental, and course dimensions interact to support BL success, offering a systematic foundation for evidence-based educational strategy prioritisation and policymaking in HEIs.

6. Conclusion

This study employed the integrated Grey-DEMATEL-ISM-MICMAC approach to systematically model the causal interdependencies, hierarchical structure, and driving-dependence classification of BL critical success factors in HEIs. The integrated analysis visualised how these factors interact within the BL system, offering a holistic understanding of their structural and functional relationships. The results identified policy support, system quality, and technical support as key causal drivers that form the structural foundation of the BL system and act as independent factors with strong driving power in achieving BL success. Course design functions as the pivotal mediator linking institutional enablers and learner outcomes, characterized by the highest centrality within the BL system. While attitude, perceived usefulness, and interaction were classified as dependent factors, representing the ultimate manifestations of system performance. These findings also indicate that well-established institutional and technological drivers do not, in themselves, guarantee positive learning experiences. In practice, the gap often arises at the level of instructional implementation, where course design shapes whether institutional investments transform into meaningful interaction, perceived value, and positive learning attitudes. These structured relationships collectively form a self-reinforcing feedback mechanism, wherein institutional policies, technical support, and system functions enable high-quality course design, which subsequently enhance learners' perceptions, engagement, and attitudes. This cyclical mechanism continuously drives improvement and innovation in BL environments.

The findings provide data-driven insights for HEI policymakers and educators, emphasizing the need to strategically allocate resources towards those strategic leverage points for optimising BL successful implementation. Specifically, institutions should prioritise institutional policy alignment, high-quality system infrastructure, strong technical support, and task-technology alignment in course design as strategic enablers that enhance learning engagement and effectiveness, ensuring the long-term sustainability of BL implementation in HEIs.

This study has several limitations that should be acknowledged. First, the participating experts were drawn from a single HEI, which may limit the generalisability of the findings, as the results may partly reflect contextual characteristics specific to this institutional environment. Second, the analysis relied primarily on expert judgement rather than large-scale empirical data. Although expert-based evaluations are widely used in systems-oriented approaches to identify structural relationships among factors, further empirical validation would strengthen the robustness of the findings. Future research could therefore extend this work through cross-cultural comparative studies to examine whether the identified causal relationships and hierarchical structures vary across different higher education contexts. In addition, longitudinal studies could further refine understanding of the dynamic interplay among BL success factors as these relationships evolve over time with institutions' growing experience in BL implementation and digital infrastructure development within technology-mediated learning environments.

Acknowledgements

This research is supported by Chiang Mai University Presidential Scholarship. We would like to express gratitude and appreciation to the experts who participate in this study for their generous sharing.

AI Statement: ChatGPT was used solely to improve language clarity and translation. All study design, analysis, and conclusions were conducted independently by the researchers.

Ethics Statement: All participation was voluntary, with informed consent obtained prior to the survey. The study complied with the Declaration of Helsinki and was approved by the Chiang Mai University Ethics Committee (COA No. 158-67).

References

Acosta-Gonzaga, E. and Ramirez-Arellano, A. 2021. The influence of motivation, emotions, cognition, and metacognition on students' learning performance: A comparative study in higher education in blended and traditional contexts. *Sage Open*, 11(2). https://doi.org/10.1177/21582440211027561

Al-Mekhlafi, A.G., Zaneldin, E., Ahmed, W., Kazim, H.Y. and Jadhav, M.D. 2025. The effectiveness of using blended learning in higher education: Students' perception. *Cogent Education*, 12(1), pp.2455228. https://doi.org/10.1080/2331186X.2025.2455228

Al-Samarraie, H. and Saeed, N. 2018. A systematic review of cloud computing tools for collaborative learning: Opportunities and challenges to the blended-learning environment. *Computers & Education*, 124, pp.77-91. https://doi.org/10.1016/j.compedu.2018.05.016

Alkarney, W. and Albraithen, M. 2018. Are critical success factors always valid for any case? A contextual perspective. *Ieee Access*, 6, pp.63496-63512. https://doi.org/10.1109/ACCESS.2018.2876792

Almerino, P., Sacro, M., Almerino, J.T., Esmoso, A.-D., Maturan, F., Atibing, N.M., Evangelista, S.S., Aro, J.L. and Ocampo, L. 2024. Interpretive structural modelling and MICMAC analysis for rethinking special mathematics education. *International Journal of Knowledge and Systems Science*, 15(1). https://doi.org/10.4018/IJKSS.353299

Alomari, M.M., El-Kanj, H., Alshdaifat, N.I. and Topal, A. 2020. A framework for the impact of human factors on the effectiveness of learning management systems. *Ieee Access*, 8, pp.23542-23558. https://doi.org/10.1109/ACCESS.2020.2970278

Alqahtani, A.Y. and Rajkhan, A.A. 2020. E-learning critical success factors during the COVID-19 pandemic: A comprehensive analysis of e-learning managerial perspectives. *Education Sciences*, 10(9), pp.216. https://doi.org/10.3390/educsci10090216

Anthony, B., Kamaludin, A., Romli, A., Raffei, A.F.M., Nincarean A/L Eh Phon, D., Abdullah, A., Ming, G.L., Shukor, N.A., Nordin, M.S. and Baba, S. 2019. Exploring the role of blended learning for teaching and learning effectiveness in institutions of higher learning: An empirical investigation. *Education and Information Technologies*, 24, pp.3433-3466. https://doi.org/10.1007/s10639-019-09941-z

Anthony, B., Kamaludin, A., Romli, A., Raffei, A.F.M., Phon, D.N.A.E., Abdullah, A. and Ming, G.L. 2022. Blended learning adoption and implementation in higher education: A theoretical and systematic review. *Technology, Knowledge and Learning*, pp.1-48. https://doi.org/10.1007/s10758-020-09477-z

Aria, A., Jafari, P. and Behifar, M. 2024. Identification of factors affecting student academic burnout in online education during the COVID-19 pandemic using grey Delphi and grey-DEMATEL techniques. *Scientific Reports*, 14(1), pp.3989. https://doi.org/10.1038/s41598-024-53233-7

Arjanto, P. and Telussa, R.P. 2024. Educational management strategies: Linking infrastructure, student activities, and academic performance. *JPPI (Jurnal Penelitian Pendidikan Indonesia)*, 10(3), pp.163-174. https://doi.org/10.29210/020244097

Asadi, S., Allison, J., Iranmanesh, M., Fathi, M., Safaei, M. and Saeed, F. 2024. Determinants of intention to use simulation-based learning in computers and networking courses: An ISM and MICMAC analysis. *IEEE Transactions on Engineering Management*, 71, pp.6015-6030. https://doi.org/10.1109/TEM.2024.3374517

Bagherian, A., Gershon, M., Kumar, S. and Kumar Mishra, M. 2024. Analyzing the relationship between digitalization and energy sustainability: A comprehensive ISM-MICMAC and DEMATEL approach. *Expert systems with Applications*, 236, pp.121193. https://doi.org/10.1016/j.eswa.2023.121193

Bai, C. and Sarkis, J. 2013. A grey-based DEMATEL model for evaluating business process management critical success factors. *International Journal of Production Economics*, 146(1), pp.281-292. https://doi.org/10.1016/j.ijpe.2013.07.011

Bashir, H. and Ojiako, U. 2020. An integrated ISM-MICMAC approach for modelling and analysing dependencies among engineering parameters in the early design phase. *Journal of Engineering Design*, 31(8-9), pp.461-483. https://doi.org/10.1080/09544828.2020.1817347

Bokolo Jr, A., Kamaludin, A., Romli, A., Mat Raffei, A.F., A/L Eh Phon, D.N., Abdullah, A., Leong Ming, G., A. Shukor, N., Shukri Nordin, M. and Baba, S. 2020. A managerial perspective on institutions' administration readiness to diffuse blended learning in higher education: Concept and evidence. *Journal of Research on Technology in Education*, 52(1), pp.37-64. https://doi.org/10.1080/15391523.2019.1675203

Castro, R. 2019. Blended learning in higher education: Trends and capabilities. *Education and Information Technologies*, 24(4), pp.2523-2546. https://doi.org/10.1007/s10639-019-09886-3

Chen, B., Chen, Y., Sun, Y., Tong, Y. and Liu, L. 2024. The measurement, level, and influence of resource allocation efficiency in universities: Empirical evidence from 13 "double first class" universities in China. *Humanities and Social Sciences Communications*, 11(1), pp.1-16. https://doi.org/10.1057/s41599-024-03461-z

Chen, L.-L. 2022. Designing online discussion for HyFlex learning. *International Journal of Educational Methodology*, 8(1), pp.191-198. https://doi.org/10.12973/ijem.8.1.191

Chuaphun, P. and Samanchuen, T. 2024. Exploring success factors and relationships in virtual learning using ISM and fuzzy MICMAC analysis. *Heliyon*, 10(7), pp.e28100. https://doi.org/10.1016/j.heliyon.2024.e28100

Dakduk, S., Santalla-Banderali, Z. and Van Der Woude, D. 2018. Acceptance of blended learning in executive education. *Sage Open*, 8(3). https://doi.org/10.1177/2158244018800647

Dalvi-Esfahani, M., Niknafs, A., Kuss, D.J., Nilashi, M. and Afrough, S. 2019. Social media addiction: Applying the DEMATEL approach. *Telematics and Informatics*, 43, pp.101250. https://doi.org/10.1016/j.tele.2019.101250

Deepu, T. and Ravi, V. 2021. Exploring critical success factors influencing adoption of digital twin and physical internet in electronics industry using grey-DEMATEL approach. *Digital Business*, 1(2). https://doi.org/10.1016/j.digbus.2021.100009

Du, Y.-W. and Shen, X.-L. 2024. Large-scale group hierarchical DEMATEL method with automatic consensus reaching. *Information Fusion*, 108, pp.102411. https://doi.org/10.1016/j.inffus.2024.102411

Du, Y. and He, H. 2025. Evaluating the multi-dimensional values of bridge-tourism integration: An empirical study using the DEMATEL—ISM—MICMAC method. *Frontiers in Public Health*, 13. https://doi.org/10.3389/fpubh.2025.1566420

Dziuban, C., Hartman, J., Cavanagh, T.B. and Moskal, P.D. 2011. Blended courses as drivers of institutional transformation. *Blended learning across disciplines: Models for implementation*. IGI Global. pp. 17-37. https://doi.org/10.4018/978-1-60960-479-0.ch002

Feng, L., He, L. and Ding, J. 2023. The association between perceived teacher support, students' ICT self-efficacy, and online English academic engagement in the blended learning context. *Sustainability*, 15(8), pp.6839. https://doi.org/10.3390/su15086839

Feng, X., Li, E., Li, J. and Wei, C. 2024. Critical influencing factors of employees' green behavior: Three-stage hybrid fuzzy DEMATEL–ISM–MICMAC approach. *Environment, development and sustainability*, 26(7), pp.17783-17811. https://doi.org/10.1007/s10668-023-03364-0

Galvis, Á.H. 2018. Supporting decision-making processes on blended learning in higher education: Literature and good practices review. *International journal of educational technology in higher education*, 15(1), pp.1-38. https://doi.org/10.1186/s41239-018-0106-1

Ghazal, S., Al-Samarraie, H. and Aldowah, H. 2018. "I am still learning": Modeling LMS critical success factors for promoting students' experience and satisfaction in a blended learning environment. *Ieee Access*, 6, pp.77179-77201. https://doi.org/10.1109/ACCESS.2018.2879677

Hair, J.F., Hult, G.T.M., Ringle, C.M., Sarstedt, M., Danks, N.P. and Ray, S. 2021. An Introduction to Structural Equation Modeling. *In* J. F. Hair Jr et al. eds. *Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R: A Workbook*. Cham, Springer International Publishing. pp. 1-29. https://doi.org/10.1007/978-3-030-80519-7_1

Hu, A., Li, Y., Li, H. and Wang, B. 2024. A comprehensive DEMATEL-ISM model-based safety analysis of the Lianghekou earth-rock dam. *Heliyon*, 10(12), pp.e33215. https://doi.org/10.1016/j.heliyon.2024.e33215

Hua, M., Wang, L. and Li, J. 2024. The impact of self-directed learning experience and course experience on learning satisfaction of university students in blended learning environments: The mediating role of deep and surface learning approach. *Frontiers in Psychology*, 14. https://doi.org/10.3389/fpsyg.2023.1278827

Huang, M., Kuang, F. and Ling, Y. 2022. EFL learners' engagement in different activities of blended learning environment. *Asian-Pacific Journal of Second and Foreign Language Education*, 7(1), pp.9. https://doi.org/10.1186/s40862-022-00136-7

Jannat Nipa, N. and Hoque, M.R. 2025. Investigating the enabling factors for effective blended learning in Bangladesh. *Information Development*, 0(0), pp.02666669251320619. https://doi.org/10.1177/02666669251320619

Javan Jafari Bojnordi, A., Zahedian Nezhad, M., Bagheri, R., Bazrafshan, M. and Sohrabi, B. 2025. Identifying, ranking and analyzing obstacles to big data analytics implementation in the healthcare industry using an ISM approach. *Discover Health Systems*, 4(1), pp.32. https://doi.org/10.1007/s44250-025-00204-y

Katsarou, E. 2021. The effects of computer anxiety and self-efficacy on L2 learners' self-perceived digital competence and satisfaction in higher education. *Journal of Education and E-Learning Research*, 8(2), pp.158-172. https://doi.org/10.20448/journal.509.2021.82.158.172

Khan, S.A., Kusi-Sarpong, S., Gupta, H., Arhin, F.K., Lawal, J.N. and Hassan, S.M. 2024. Critical factors of digital supply chains for organizational performance improvement. *IEEE Transactions on Engineering Management*, 71, pp.13727-13741. https://doi.org/10.1109/TEM.2021.3052239

Konstantinidou, A. and Nisiforou, E. 2022. Assuring the quality of online learning in higher education: Adaptations in design and implementation. *Australasian Journal of Educational Technology*, 38(4), pp.127-142. https://doi.org/10.14742/ajet.7910

Lan, Z., Pau, K., Mohd Yusof, H. and Huang, X. 2022. Hierarchical topological model of the factors influencing adolescents' non-suicidal self-injury behavior based on the DEMATEL-TAISM method. *Sci Rep*, 12(1), pp.17238. https://doi.org/10.1038/s41598-022-21377-z

Laohajaratsang, T. 2017. A case study of how the technological infrastructure of Chiang Mai University supports blended learning practices. *In* C. P. Lim and L. Wang eds. *Blended learning for quality higher education: Selected case studies on implementation from Asia-Pacific*. France, UNESCO. pp. 127-152. https://unesdoc.unesco.org/ark:/48223/pf0000246851

Law, K.M., Geng, S. and Li, T. 2019. Student enrollment, motivation and learning performance in a blended learning environment: The mediating effects of social, teaching, and cognitive presence. *Computers & Education*, 136, pp.1-12. https://doi.org/10.1016/j.compedu.2019.02.021

Li, J. and Xiao, Y. 2024. Analysis of influencing factors on review efficiency of multidisciplinary scientific research projects using DEMATEL with a 5-point scale. *PloS one*, 19(12), pp.e0315349. https://doi.org/10.1371/journal.pone.0315349

Li, X. and Zhu, W. 2022. System quality, information quality, satisfaction and acceptance of online learning platform among college students in the context of online learning and blended learning. *Frontiers in Psychology*, 13, pp.1054691. https://doi.org/10.3389/fpsyg.2022.1054691

Liu, W., Hu, Y. and Huang, Q. 2024. Research on critical factors influencing organizational resilience of major transportation infrastructure projects: A hybrid Fuzzy DEMATEL-ISM-MICMAC approach. *Buildings*, 14(6), pp.1598. https://doi.org/10.3390/buildings14061598

Liu, X., Deng, Q., Gong, G., Zhao, X. and Li, K. 2021. Evaluating the interactions of multi-dimensional value for sustainable product-service system with grey DEMATEL-ANP approach. *Journal of Manufacturing Systems*, 60, pp.449-458. https://doi.org/10.1016/j.jmsy.2021.07.006

Liu, X. and Yodmongkol, P. 2023. Influencing factors of blended learning in higher education: A systematic literature review. 2023 International Conference on University Teaching and Learning (InCULT), Shah Alam, Malaysia, IEEE. https://doi.org/10.1109/InCULT59088.2023.10482667

Liu, Y. 2025. Five perspectives on the digital: A sociological interpretation. *The Journal of Chinese Sociology*, 12(1), pp.21. https://doi.org/10.1186/s40711-025-00248-w

Lu, M.-T., Huang, J.-C. and Wang, M.-H. 2024. Evaluating the mobile learning quality for promoting the user needs in Taiwanese higher education during COVID-19. *International Journal of Information Technology & Decision Making*, 23(02), pp.629-649. https://doi.org/10.1142/S0219622023500232

Majeed, M. and Rehan Dar, F. 2022. Investigating the efficacy of blended learning in ESL classrooms. *Cogent Education*, 9(1), pp.2133500. https://doi.org/10.1080/2331186X.2022.2133500

McCarthy, S. and Palmer, E. 2023. Defining an effective approach to blended learning in higher education: A systematic review. *Australasian Journal of Educational Technology*, 39(2), pp.98-114. https://doi.org/10.14742/ajet.8489

Mielikäinen, M. 2022. Towards blended learning: Stakeholders' perspectives on a project-based integrated curriculum in ICT engineering education. *Industry and Higher Education*, 36(1), pp.74-85. https://doi.org/10.1177/0950422221994471

Min, W. and Yu, Z. 2023. A systematic review of critical success factors in blended learning. *Education Sciences*, 13(5), pp.469. https://doi.org/10.3390/educsci13050469

Mizza, D., Reese, M. and Malouche, D. 2025. Flipped classroom evaluation and blended learning potential: A case study of engagement and inclusion in quantitative education. *Smart Learning Environments*, 12(1), pp.56. https://doi.org/10.1186/s40561-025-00412-2

Mohammadi, M., Paasivara, M. and Kasurinen, J. 2025. Blended learning in higher education: Good practices in platforms and teachers support, enhancing students motivation. *Education and Information Technologies*. https://doi.org/10.1007/s10639-025-13770-8

Nikou, S. and Maslov, I. 2023. Finnish university students' satisfaction with e-learning outcomes during the COVID-19 pandemic. *International Journal of Educational Management*, 37(1), pp.1-21. https://doi.org/10.1108/IJEM-04-2022-0166

Nortvig, A.-M., Petersen, A.K. and Balle, S.H. 2018. A literature review of the factors influencing e-learning and blended learning in relation to learning outcome, student satisfaction and engagement. *Electronic journal of e-Learning*, 16(1), pp.46-55. https://academic-publishing.org/index.php/ejel/article/view/1855

Opricovic, S. and Tzeng, G.-H. 2003. Defuzzification within a multicriteria decision model. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 11(05), pp.635-652. https://doi.org/10.1142/S0218488503002387

Pattanasak, P., Anantana, T., Paphawasit, B. and Wudhikarn, R. 2022. Critical factors and performance measurement of business incubators: A systematic literature review. *Sustainability*, 14(8), pp.4610. https://doi.org/10.3390/su14084610

Pinili, L., Almerino Jr, P., Tenerife, J.J., Evangelista, S.S., Almerino, J.G., Aro, J.L., Arnaiz, V., Opingo, K.M., Deniega, J. and Revalde, H. 2024. Uncertain causality analysis of critical success factors of special education mathematics teaching. *Open Education Studies*, 6(1), pp.20220232. https://doi.org/10.1515/edu-2022-0232

Pollard, C. and Cater-Steel, A. 2009. Justifications, strategies, and critical success factors in successful ITIL omplementations in U.S. and Australian companies: An exploratory study. *Information Systems Management*, 26(2), pp.164-175. https://doi.org/10.1080/10580530902797540

Prifti, R. 2022. Self–efficacy and student satisfaction in the context of blended learning courses. *Open Learning: The Journal of Open, Distance and e-Learning*, 37(2), pp.111-125. https://doi.org/10.1080/02680513.2020.1755642

Primadasa, R., Kusrini, E., Mansur, A. and Masudin, I. 2025. Integrating DEMATEL-ISM-MICMAC: An interconnected model of halal-sustainable supply chain management (HSSCM) indicatorsfor SMEs. *Journal of Islamic Marketing*. https://doi.org/10.1108/jima-07-2024-0303

Quiñones, R.S., Caladcad, J.A.A., Himang, C.M., Quiñones, H.G., Castro, C.J., Caballes, S.A.A., Abellana, D.P.M., Jabilles, E.M.Y. and Ocampo, L.A. 2020. Using Delphi and fuzzy DEMATEL for analyzing the intertwined relationships of the barriers of university technology transfer: Evidence from a developing economy. *International Journal of Innovation Studies*, 4(3), pp.85-104. https://doi.org/10.1016/j.ijis.2020.07.002

Raj, A. and Sah, B. 2019. Analyzing critical success factors for implementation of drones in the logistics sector using grey-DEMATEL based approach. *Computers & Industrial Engineering*, 138, pp.106118. https://doi.org/10.1016/j.cie.2019.106118

Rasheed, R.A., Kamsin, A. and Abdullah, N.A. 2020. Challenges in the online component of blended learning: A systematic review. *Computers & Education*, 144, pp.103701. https://doi.org/10.1016/j.compedu.2019.103701

Ren, W., Wang, R., Zhao, X. and Shi, J. 2025. A comparative analysis of the advantages and disadvantages of blended learning and traditional teaching in a multidisciplinary context. *Frontiers in Education*, 10. https://doi.org/10.3389/feduc.2025.1659590

Rix, R.W. 2011. Blended learning: Perspectives on mixing online and offline communities of enquiry. *E-Learning and Digital Media*, 8(4), pp.423-433. https://doi.org/10.2304/elea.2011.8.4.423

Roca, J.C., Chiu, C.-M. and Martínez, F.J. 2006. Understanding e-learning continuance intention: An extension of the Technology Acceptance Model. *International Journal of human-computer studies*, 64(8), pp.683-696. https://doi.org/10.1016/j.ijhcs.2006.01.003

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M. and Tarantola, S. 2008. *Global sensitivity analysis: The primer*: John Wiley & Sons. https://doi.org/10.1002/9780470725184

Samala, A.D., Papadakis, S. and Rawas, S. 2025. Global insights into mobile learning in higher education: A PRISMA-guided bibliometric analysis from 2007 to 2023. *International Journal of Educational Reform*, 0(0). https://doi.org/10.1177/10567879251341869

Sareen, S. and Mandal, S. 2024. Challenges of blended learning in higher education across global north-south: A systematic and integrative literature review. *Social Sciences & Humanities Open*, 10, pp.101011. https://doi.org/10.1016/j.ssaho.2024.101011

Selim, H.M. 2007. Critical success factors for e-learning acceptance: Confirmatory factor models. *Computers & Education*, 49(2), pp.396-413. https://doi.org/10.1016/j.compedu.2005.09.004

Si, S.-L., You, X.-Y., Liu, H.-C. and Zhang, P. 2018. DEMATEL technique: A systematic review of the state-of-the-art literature on methodologies and applications. *Mathematical problems in Engineering*, 2018(1), pp.3696457. https://doi.org/10.1155/2018/3696457

Su, F., Zou, D., Wang, L. and Kohnke, L. 2023. Student engagement and teaching presence in blended learning and emergency remote teaching. *Journal of Computers in Education*, pp.1-26. https://doi.org/10.1007/s40692-023-00263-1

Taghizadeh, M. and Hajhosseini, F. 2021. Investigating a blended learning environment: Contribution of attitude, interaction, and quality of teaching to satisfaction of graduate students of TEFL. *The Asia-Pacific Education Researcher*, 30, pp.459-469. https://doi.org/10.1007/s40299-020-00531-z

Thakkar, J.J. 2021. *Multi-criteria decision making*: Springer. https://doi.org/10.1007/978-981-33-4745-8

Thompson, R., Compeau, D. and Higgins, C. 2008. Intentions to use information technologies: An integrative model. *End-User Computing: Concepts, Methodologies, Tools, and Applications*. IGI Global Scientific Publishing. pp. 1469-1488. https://doi.org/10.4018/978-1-59904-945-8.ch100

Tuapawa, K. 2017. Identifying key stakeholders in blended tertiary environments: Experts' perspectives. *International Journal of Information and Communication Technology Education (IJICTE)*, 13(4), pp.40-52. https://doi.org/10.4018/IJICTE.2017100104

Verma, R., Kumar, A., Kumar, P. and Kumar, A. 2025. Exploring research dynamics in Open, Distance and Online Learning (ODOL): A scientometric study. *Discover Education*, 4(1), pp.353. https://doi.org/10.1007/s44217-025-00776-0

Vo, M.H., Zhu, C. and Diep, A.N. 2020. Students' performance in blended learning: Disciplinary difference and instructional design factors. *Journal of Computers in Education*, 7(4), pp.487-510. https://doi.org/10.1007/s40692-020-00164-7

Wang, J., Qin, Y., He, P. and Yan, W. 2025. Research on smart construction site evaluation model based on DEMATEL-ANP method. *Buildings*, 15(17), pp.3077. https://doi.org/10.3390/buildings15173077

Wang, Y., Han, X. and Yang, J. 2015. Revisiting the blended learning literature: Using a complex adaptive systems framework. *Journal of Educational Technology & Society*, 18(2), pp.380-393. https://doi.org/10.2307/jeductechsoci.18.2.380

Więckowski, J. and Sałabun, W. 2023. Sensitivity analysis approaches in multi-criteria decision analysis: A systematic review. *Applied Soft Computing*, pp.110915. https://doi.org/10.1016/j.asoc.2023.110915

Wu, J. and Liu, W. 2013. An empirical investigation of the critical factors affecting students' satisfaction in EFL blended learning. *Journal of Language Teaching and Research*, 4(1), pp.176. https://doi.org/10.4304/jltr.4.1.176-185

Wu, P., Yang, L., Hu, X., Li, B., Liu, Q., Wang, Y. and Huang, J. 2022. How K12 teachers' readiness influences their intention to implement STEM education: Exploratory study based on decomposed theory of planned behavior. *Applied Sciences*, 12(23), pp.11989. https://doi.org/10.3390/app122311989

Xia, X., Govindan, K. and Zhu, Q. 2015. Analyzing internal barriers for automotive parts remanufacturers in China using grey-DEMATEL approach. *Journal of cleaner production*, 87, pp.811-825. https://doi.org/10.1016/j.jclepro.2014.09.044

Zhang, S., Liu, J., Li, Z., Xiahou, X. and Li, Q. 2024. Analyzing critical factors influencing the quality management in smart construction site: A DEMATEL-ISM-MICMAC based approach. *Buildings*, 14(8), pp.2400. https://doi.org/10.3390/buildings14082400

Zhao, S. and Song, J. 2021. What kind of support do teachers really need in a blended learning context? *Australasian Journal of Educational Technology*, 37(4), pp.116-129. https://doi.org/10.14742/ajet.6592

Zhou, L., Tang, M. and Liu, J. 2025. Analysis of factors influencing MOOC quality based on I-DEMATEL-ISM method. *Systems and Soft Computing*, 7, pp.200220. https://doi.org/10.1016/j.sasc.2025.200220

Zhou, X., Smith, C.J.M. and Al-Samarraie, H. 2023. Digital technology adaptation and initiatives: A systematic review of teaching and learning during COVID-19. *Journal of Computing in Higher Education*, pp.1-22. https://doi.org/10.1007/s12528-023-09376-z

Generative AI and Knowledge Mapping in Programming Education: Student Learning and Engagement

Athitaya Nitchot[1] **and Lester Gilbert**[2]
[1]Creative Media and Digital Technologies Program, Prince of Songkla University International College, Prince of Songkla University, Hatyai, Thailand
[2]Electronics and Computer Science, University of Southampton, UK

athitaya.nitchot@gmail.com
lg11@soton.ac.uk

https://doi.org/10.34190/ejel.24.2.4587

Abstract: This study examines how Generative AI and knowledge-mapping tools support student learning and engagement in programming education. A quasi-experimental design was conducted with 30 undergraduate students enrolled in an object-oriented programming course, where participants used both tools across a four-week intervention. Data were collected through task performance and learner perception surveys. The results indicate that students reported higher ease of use and immediate support when using Generative AI, while knowledge mapping was associated with stronger support for conceptual understanding and reflective learning in later stages. These findings suggest that the two approaches support different aspects of learning, with Generative AI facilitating rapid clarification and knowledge-mapping tools encouraging structured conceptual engagement. The study contributes to the e-learning field by providing empirical insight into how different forms of learning support function within the same instructional context. Rather than positioning the tools as direct alternatives, the findings highlight their complementary pedagogical roles and offer guidance for integrating adaptive AI support with structured learning approaches in programming education.

Keywords: Generative AI, Knowledge mapping, Mytelemap, Programming education, Student learning outcomes

1. Introduction

Artificial intelligence (AI) continues to reshape the educational landscape by offering tools that enhance both teaching and learning experiences. Among these innovations, Generative AI technologies have emerged as powerful agents of change, capable of generating personalized content and supporting self-directed learning. Tools such as OpenAI's ChatGPT exemplify this trend by offering real-time interaction, instant feedback, and tailored study materials, which make them particularly useful in accommodating diverse learner needs. Complementing these advances, knowledge mapping tools like Mytelemap provide structured and visual approaches to learning. By enabling learners to map out concepts and relationships between topics, Mytelemap supports competence-based learning and encourages deeper engagement with content. Its design aligns well with educational goals that prioritize conceptual understanding, self-regulated learning, and critical thinking. Mytelemap is considered in this study as an example of a knowledge-mapping approach rather than as a standalone tool for evaluation. Knowledge mapping has been widely used in educational research to support conceptual understanding, schema development, and reflective learning. The inclusion of Mytelemap therefore represents a broader category of visualization-based learning support, enabling a comparison with AI-based approaches within the same instructional context. Despite their promise, both Generative AI and Mytelemap face challenges. Generative AI raises concerns about content reliability, ethical use, and academic integrity, while Mytelemap requires users to invest time and effort to master its mapping techniques. Understanding the pedagogical differences between these tools is important for educators seeking to integrate them effectively into programming instruction. As the two tools serve different pedagogical functions, this study does not treat them as directly equivalent. Instead, the comparison focuses on how different forms of learning support—Generative AI and knowledge mapping—are associated with differences in student learning processes and experiences within the same instructional context.

This study investigates how different forms of learning support—Generative AI and knowledge mapping—are associated with differences in student performance and perceptions in programming learning. It examines their perceptions of each tool's usability, learning support, and overall effectiveness. The research also investigates how each tool is associated with differences in learning performance and student satisfaction. By analyzing these outcomes, the study aims to provide practical insights into the role of emerging technologies in education and offers recommendations for integrating them into hybrid learning environments. Also, the present study is

Cite this article: Nitchot, A. and Gilbert, L. 2026. "Generative AI and Knowledge Mapping in Programming Education: Student Learning and Engagement", *Electronic Journal of e-Learning*, 24(2), pp 132-148, https://doi.org/10.34190/ejel.24.2.4587

theoretically informed by Cognitive Load Theory (Sweller, 1994) and Constructivist learning principles. From a Cognitive Load perspective, Generative AI may reduce extraneous load by offering immediate, simplified explanations that help novices progress quickly through early programming tasks. In contrast, Mytelemap's knowledge-mapping processes align with constructivist theory, which emphasizes active meaning-making, integration of new information with prior knowledge, and the construction of conceptual structures. Mapping requires learners to externalize relationships among programming concepts, thereby promoting cognitive load and deeper schema formation. Framing the comparison through these theories provides explanatory power for understanding why AI-driven tools accelerate comprehension while visualization-based tools enhance conceptual depth.

This study contributes empirical insight by examining how a Generative AI tool and a visualization-based knowledge-mapping tool (Mytelemap) support learning within the same programming course context. Whereas prior studies examined each tool independently, relatively few studies have examined how their differing learning mechanisms—AI-mediated adaptive feedback and competence-based visual structuring—may shape student outcomes within the same instructional context. The findings contribute to ongoing discussions in technology-enhanced learning by providing insight into when each tool may be pedagogically advantageous, thereby informing the design of hybrid learning models that integrate AI-driven personalization with structured conceptual mapping.

Despite the increasing use of both Generative AI and knowledge-mapping tools in education, the research problem remains insufficiently articulated in existing literature. Prior studies have typically examined these technologies separately, with limited attention to how they function within the same learning context. As a result, there is still a lack of clarity regarding how different forms of learning support—AI-driven adaptive feedback and visualization-based knowledge structuring—affect student learning processes and outcomes in programming education.

Based on the above considerations, this study addresses the following research questions:

RQ1: *How does the use of Generative AI compare with Mytelemap in supporting students' programming task performance?*

RQ2: *How do students perceive the usability and learning support provided by Generative AI and Mytelemap?*

RQ3: *In what ways do the two tools differ in supporting conceptual understanding and engagement in programming learning?*

The analytical focus of this study is to examine how these two forms of learning support differ in their underlying learning mechanisms and how these differences are reflected in student performance and perceptions.

Section 2 provides an overview of Generative AI and Mytelemap, examining their roles in education and summarizing previous comparative studies that highlight their respective advantages and limitations. Section 3 describes the research methodology employed in this study, including the design of the experiment, data collection procedures, and the evaluation criteria used to assess the tools' effectiveness. Section 4 presents the results and discusses key findings, focusing on how each tool impacted students' learning outcomes and user satisfaction. Finally, Section 5 concludes the study by summarizing the main contributions, addressing the limitations, and offering recommendations for future research, particularly on the integration of Generative AI with knowledge mapping technologies to create more effective hybrid learning solutions.

2. Generative AI and Mytelemap

2.1 Generative AI and its Uses in Education

Generative Artificial Intelligence (Generative AI) is an emerging technology that is increasingly influencing educational practices. Its capabilities are influencing how learning content is created and delivered to students, and in turn how students process and respond to learning tasks, offering new opportunities for engagement and efficiency. Generative AI refers to systems that leverage machine learning techniques, particularly deep learning models such as large language models (LLMs), to produce original content in response to user inputs (Bender et al., 2021). These outputs can include natural language text, images, audio, video, or even software code, and are typically generated in a way that closely mimics human reasoning and creativity (Bender et al., 2021). In the educational context, Generative AI offers several potential benefits for both learners and educators, although its adoption comes with challenges that require careful consideration.

One of the most significant contributions of Generative AI in education is its ability to support personalized learning. Traditional educational models often struggle to accommodate diverse learner needs, particularly in large classrooms where individual attention is limited. Generative AI tools can provide customized learning experiences aligned with students' knowledge level and pace (Khreisat et al., 2024). These tools are capable of generating personalized quizzes, flashcards, study guides, and summaries that align with individual learner goals. For instance, Xodabande, Atai and Hashemi (2024) found that university students who used AI-generated digital flashcards demonstrated significantly higher retention of technical vocabulary compared to those relying on conventional learning materials. Such adaptive learning support facilitates differentiated instruction, ensuring that learners with varying capabilities and backgrounds can receive content that meets their specific needs.

Generative AI also enhances the interactivity of the learning experience. AI-powered virtual tutors can simulate personalized, one-on-one instruction by responding to learners' questions in real time. These systems, designed to emulate the guidance of a human teacher, provide detailed explanations, step-by-step problem-solving assistance, and contextually relevant examples to reinforce comprehension. Were (2022) highlights the potential of these AI-driven tutoring systems to bridge educational gaps, particularly in under-resourced environments where access to qualified instructors is limited. By providing immediate, adaptive feedback, Generative AI enables students to clarify doubts without delay, promoting continuous learning and reducing dependence on external instruction.

In addition to enhancing learner engagement, Generative AI streamlines various administrative and instructional processes for educators. One of its practical applications is automating the generation and grading of assessments. By leveraging natural language processing and machine learning algorithms, AI systems can automatically generate exam questions, assignments, and project guidelines tailored to specific course objectives (Eden, Chisom and Adeniyi, 2024). Furthermore, AI-powered grading systems offer immediate and consistent feedback on student submissions, freeing educators from time-intensive evaluation tasks and allowing them to focus on more meaningful interactions with students (Saqr, Al-Somali and Sarhan, 2024). Automated feedback systems also enable students to track their performance in real time, identifying areas for improvement and adapting their learning strategies accordingly.

Beyond classroom applications, Generative AI contributes to developing educational resources and materials. Tools like ChatGPT can assist instructors in designing lesson plans, writing lecture notes, and creating instructional videos or interactive modules (Hutson et al., 2023). This capacity to generate high-quality educational content reduces the preparation time required by educators, allowing them to dedicate more attention to facilitating discussions, mentoring students, and fostering collaborative learning environments.

Despite these benefits, the integration of Generative AI into education raises several challenges and ethical concerns. One of the most pressing issues relates to the accuracy and reliability of AI-generated content. While Generative AI models are capable of producing plausible and fluent text, they are also prone to generating factually incorrect or misleading information, a phenomenon commonly referred to as "hallucination" (Bender et al., 2021). This can have serious consequences in educational settings, where students may unknowingly rely on erroneous explanations or data. Ensuring that AI-generated materials are accurate and verifiable is a significant challenge that requires the development of robust validation frameworks and critical media literacy skills among learners (Singh et al., 2025).

Academic integrity is another area of concern. The ease with which Generative AI tools can produce essays, reports, computer code, and other academic work has led to increased instances of plagiarism and unauthorized assistance (Hutson et al., 2023). As students gain access to tools capable of completing assignments with minimal effort, educators face difficulties in assessing authentic student learning and ensuring fair evaluation practices. To address these challenges, institutions are adopting AI-detection tools, revising assessment methods, and encouraging the development of original, process-based assignments that emphasize critical thinking and personal reflection (Singh et al., 2025).

Data privacy and security are additional ethical considerations associated with the use of Generative AI in education. Many AI-powered platforms require access to user data to personalize learning experiences and improve model performance. However, this data collection raises concerns about user consent, data protection, and potential breaches (Khreisat et al., 2024). Ensuring compliance with data privacy regulations such as the General Data Protection Regulation (GDPR) and establishing transparent data governance policies are essential for protecting learners' sensitive information and building trust in AI systems.

Furthermore, Generative AI systems may inadvertently perpetuate biases present in their training data. Because LLMs are trained on vast datasets collected from the internet, they can reproduce and amplify existing social, cultural, and ideological biases (Bender et al., 2021). This can lead to discriminatory or biased outputs, which undermine inclusivity and equity in education. Addressing these issues requires careful curation of training datasets, regular auditing of AI systems, and the implementation of fairness-aware algorithms designed to minimize biased outcomes (Eden, Chisom and Adeniyi, 2024).

Despite these challenges, proponents argue that Generative AI has the potential to advance inclusive and equitable education when implemented responsibly. AI's ability to deliver scalable, personalized instruction can help close educational gaps, particularly for marginalized and underserved populations (Were, 2022). For example, AI-powered translation services and multilingual learning resources can support students who speak different languages, while adaptive learning tools can assist students with disabilities by offering customized content and accessibility features (Strielkowski et al., 2024). These inclusive capabilities highlight the transformative potential of Generative AI in addressing long-standing disparities in educational access and quality.

Another important consideration is the role of Generative AI in fostering lifelong learning. In today's rapidly evolving knowledge economy, learners must continuously update their skills and knowledge to remain competitive. Generative AI supports lifelong learning by providing on-demand access to up-to-date information, personalized learning pathways, and micro-learning opportunities tailored to individual interests and career goals (Saqr, Al-Somali and Sarhan, 2024). This flexibility empowers learners to take control of their educational journeys and engage in self-directed learning at their own pace.

Generative AI offers a range of applications in education. Its ability to personalize learning, facilitate engagement, automate instructional tasks, and provide scalable support holds promise for enhancing learning outcomes and improving educational access. However, the successful integration of Generative AI into education requires a balanced approach that addresses concerns related to content accuracy, academic integrity, data privacy, bias, and ethical use. As institutions and educators explore the adoption of Generative AI tools, it is essential to develop guidelines and best practices that ensure these technologies are used responsibly and effectively to support meaningful learning experiences.

Future research should focus on evaluating the long-term impacts of Generative AI in education, particularly in terms of student learning outcomes, teacher roles, and curriculum design. Studies exploring hybrid approaches that combine AI-driven personalization with human-centered instruction may offer insights into creating balanced and effective learning environments. Additionally, interdisciplinary collaboration among educators, technologists, ethicists, and policymakers is needed to establish comprehensive frameworks for the ethical and pedagogical use of Generative AI in education.

2.2 Overview of Mytelemap in Education

Mytelemap, a digital tool designed for knowledge mapping, has gained traction in educational contexts for its ability to visually organize and connect learners' knowledge (Nitchot, Wettayaprasit and Gilbert, 2019). Knowledge mapping, a pedagogical technique, helps students visualize the knowledge "landscape" of their learning. Mytelemap's intuitive interface and dynamic visualization capabilities make it a valuable asset in both traditional and online learning environments. The Mytelemap prototype application (www.mytelemap.com) is a web-based tool that allows for the display of knowledge mapping. Clicking on the visualization allows viewers to interact with it, and the screenshot shown illustrates how the mapping overview is arranged (Figure 1). The algorithm finds the connections based on the subject matter keywords when students click on a mapping node. Mytelemap filters the potential web links by selecting mapping node that corresponds to each subject area. Learners can discover their competencies, spot inconsistencies, integrate and adapt as needed, and test their new knowledge through additional map development and traversal through the traversal of a pedagogical knowledge mapping.

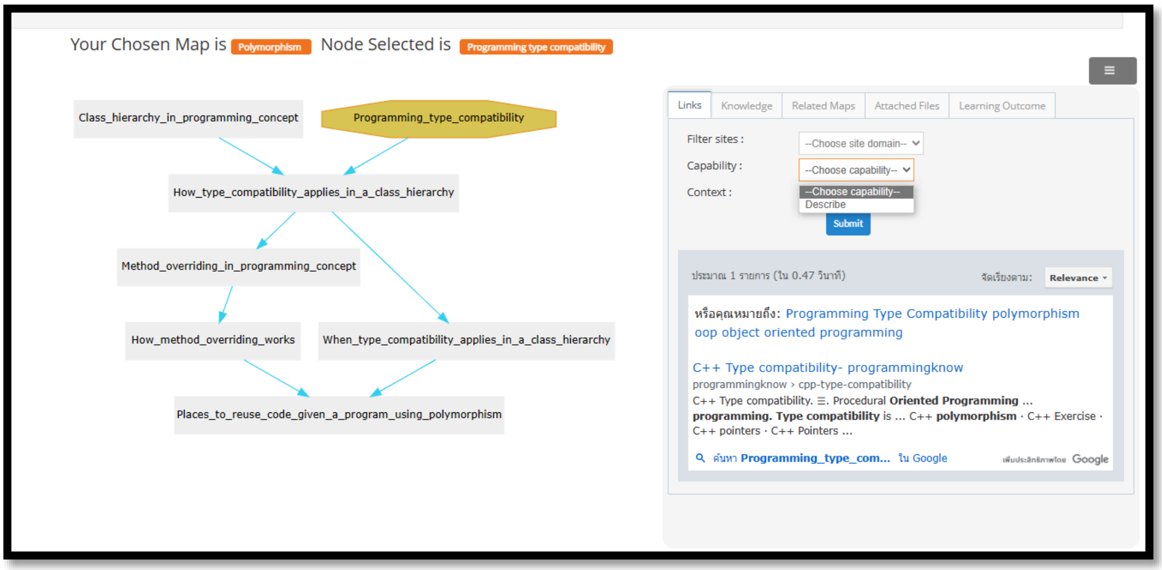


Figure 1: Mytelemap Prototype Showing the Knowledge Map of the Topic 'Polymorphism'

One of the tool's primary innovations lies in its support for competence mapping. Competence-based education requires learners to develop and demonstrate specific skills and knowledge in a structured, transparent manner. Mytelemap enables users to create maps that reflect their competencies, explicitly outlining the knowledge and skills they have acquired as well as those they still need to develop. In the studies related to the uses of Mytelemap, Nitchot and Gilbert (2024a) demonstrated that Mytelemap effectively supports competence articulation, offering learners a clearer picture of their learning trajectories and providing them with personalized recommendations for further study. The ability to track one's competencies over time fosters self-regulated learning, as learners are empowered to take responsibility for their educational progress. Moreover, Mytelemap enhances motivation and learner engagement. In a recent study on programming courses, Nitchot and Gilbert (2024b) found that students using Mytelemap reported increased motivation and satisfaction. By visualizing complex programming concepts, such as inheritance and polymorphism, students gained greater clarity and confidence in their understanding. The competence mapping functionality allowed them to see their learning progress in real time, which fostered a sense of achievement and encouraged continued engagement with the material.

Another significant advantage of Mytelemap is its facilitation of collaborative learning. The platform allows multiple users to co-create and share knowledge maps, promoting peer collaboration and collective knowledge construction. As Rahayu, Ferdiana and Kusumawardani (2022) observed, collaborative knowledge mapping fosters critical thinking and deepens conceptual understanding by encouraging learners to negotiate meanings and integrate diverse perspectives into a unified map. Mytelemap's collaborative features support synchronous and asynchronous learning, making it a valuable tool for group projects and blended learning environments. In addition to supporting individual and collaborative learning, Mytelemap also plays a crucial role in reflective practice. Nitchot and Gilbert (2025) conducted a case study in which students used Mytelemap to map out competencies required for constructing a holographic projector. The study found that learners benefited from the opportunity to reflect on their existing competencies and identify areas for improvement. By explicitly visualizing the links between acquired and required skills, Mytelemap encouraged metacognitive awareness and self-directed learning, enabling students to make informed decisions about their learning priorities.

Mytelemap represents a significant advancement in knowledge mapping tools for education. Its pedagogically informed design, emphasis on competence-based learning, and support for personalized and collaborative learning make it a valuable asset for both learners and educators. The growing body of research, particularly by (Nitchot and Gilbert, 2024a; Nitchot and Gilbert, 2024b; Nitchot and Gilbert, 2025), demonstrates Mytelemap's effectiveness in promoting critical thinking, motivation, and self-regulated learning. As educational institutions seek to harness the power of knowledge mapping, Mytelemap offers a promising solution that aligns with contemporary educational goals. Nevertheless, its future development—especially with the integration of AI—must address challenges related to usability, accessibility, and ethical considerations to ensure its impact remains positive and equitable across diverse learning environments.

2.3 Comparative Studies between Generative AI and Non-AI Tools

The rapid development of educational technology has introduced a variety of tools aimed at improving learning experiences and outcomes. Among these, Generative Artificial Intelligence (Generative AI) systems and Non-AI knowledge mapping tools such as Mytelemap represent two distinct learning support mechanisms with different pedagogical roles. While both aim to support learning and knowledge acquisition, they operate on fundamentally different principles. Generative AI leverages machine learning algorithms to generate adaptive content and provide personalized feedback, whereas Mytelemap focuses on helping students visually structure and organize knowledge. Although existing research explores the individual use of Generative AI and knowledge mapping tools in education, no prior studies have directly compared Generative AI and Mytelemap in terms of their educational effectiveness. This section synthesizes available literature on each tool type and presents comparative insights based on recent experimental findings.

Generative AI tools such as ChatGPT and similar systems have been widely recognized for their ability to deliver personalized learning experiences. These tools can generate instant explanations, summaries, quizzes, and examples tailored to individual learners' needs (Khreisat et al., 2024). Their interactive, conversational nature allows learners to ask questions and receive immediate responses, which can help clarify complex concepts and support self-paced learning (Saqr, Al-Somali and Sarhan, 2024). Studies have shown that students who use Generative AI tools often experience increased motivation and engagement due to the immediacy and personalization of the feedback they receive (Ward et al., 2025). Furthermore, Generative AI has been noted for its potential to support inclusive learning by adapting content to accommodate diverse learning styles (Eden, Chisom and Adeniyi, 2024).

In contrast, Mytelemap is a Non-AI knowledge mapping tool that assists learners in visually organizing and connecting ideas within a given domain. Knowledge mapping, as a pedagogical strategy, encourages students to structure information hierarchically and explore the relationships between concepts (Nitchot, Wettayaprasit and Gilbert, 2019). Mytelemap facilitates this process through an intuitive interface that allows users to create nodes and links, enabling them to develop a comprehensive representation of their understanding of a subject. This process promotes deeper learning by requiring students to actively engage with the material, evaluate connections, and reflect on how new information fits within their existing knowledge structures (Rahayu, Ferdiana and Kusumawardani, 2022). By fostering this metacognitive engagement, tools like Mytelemap support long-term retention and conceptual understanding.

Although no previous studies have directly compared Generative AI tools and Mytelemap, recent experimental research provides insight into how these tools perform relative to one another in specific educational contexts. In an object-oriented programming course, students used both tools to support their learning. Generative AI was found to be more effective in terms of ease of use and providing quick, adaptive responses. Many students appreciated the AI tool's ability to generate immediate answers and examples, which they found particularly helpful during early-stage learning when familiarity with programming concepts was still developing. These findings are consistent with the broader literature, which highlights Generative AI's ability to facilitate quick access to information and provide tailored support (Saqr, Al-Somali and Sarhan, 2024; Khreisat et al., 2024).

However, as students progressed through more complex topics such as polymorphism and inheritance, they encountered limitations in Generative AI's content accuracy and reliability. While Generative AI tools like ChatGPT can generate helpful explanations, they are not immune to producing incorrect or oversimplified information (Bender et al., 2021). Some students in the study expressed concerns about the need to cross-check AI-generated content with authoritative sources, which could hinder independent learning if not managed carefully.

On the other hand, Mytelemap was found to be more effective in promoting deep conceptual understanding, particularly as students engaged with advanced programming topics. Students using Mytelemap demonstrated improved abilities to visualize relationships between concepts, such as how inheritance relates to encapsulation and polymorphism. The knowledge mapping process required them to actively organize information and make decisions about concept relationships, leading to a more comprehensive understanding of object-oriented programming. These findings align with existing research on the effectiveness of knowledge mapping tools in fostering critical thinking and deep learning (Rahayu, Ferdiana and Kusumawardani, 2022; Nitchot, Wettayaprasit and Gilbert, 2019).

One of the primary differences observed between Generative AI and Mytelemap was in the level of cognitive engagement required from learners. Generative AI tools offered convenience and efficiency, which made them attractive for quick reference and immediate feedback. However, this convenience sometimes led to passive consumption of information rather than active learning. Some students became overly reliant on AI-generated answers without critically evaluating them, a concern echoed in the literature regarding potential over-reliance on AI (Strielkowski et al., 2024; Singh et al., 2025). In contrast, Mytelemap demanded more active involvement from students, as they had to build their knowledge maps manually. This process was initially perceived as more time-consuming and complex, but over time, students reported gaining a deeper understanding of the material. The need to explicitly identify and connect concepts encouraged reflective learning and a more nuanced understanding of the course content. While the tool presented an initial learning curve, particularly for students unfamiliar with knowledge mapping strategies, many participants expressed satisfaction with Mytelemap's ability to enhance their conceptual clarity once they became comfortable with its use.

Usability and accessibility were also key areas of comparison. Generative AI tools generally require little to no training, making them highly accessible to learners at all levels of technical proficiency (Eden, Chisom and Adeniyi, 2024). Their natural language interfaces allow users to interact with the system easily, reducing the cognitive load associated with learning how to use the tool itself. In contrast, Mytelemap's interface, while designed to be intuitive, still requires users to invest time in learning how to create effective knowledge maps. This additional effort can act as a barrier for some learners, though it also provides opportunities for skill development in information organization and critical thinking. Ethical considerations further distinguish these two tools. Generative AI tools typically rely on large datasets and user interactions to refine their outputs, raising concerns about data privacy and informed consent (Khreisat et al., 2024). Additionally, the ease with which students can generate essays, code, or other assignments using AI increases the risk of academic dishonesty (Hutson et al., 2023). In contrast, Non-AI tools like Mytelemap generally operate within controlled educational environments where data privacy concerns are easier to manage. Because Mytelemap emphasizes active participation, it also reduces opportunities for students to engage in plagiarism or passive learning.

Despite the differences between Generative AI and Mytelemap, there is potential for these tools to complement each other when integrated thoughtfully into educational settings. For instance, Generative AI could assist in generating preliminary content or suggesting resources that students could then organize and refine within a knowledge mapping tool like Mytelemap. Such hybrid approaches have been proposed in the literature as a way to balance the efficiency and personalization of AI with the deep learning and critical thinking encouraged by traditional educational tools (Hutson et al., 2023; Eden, Chisom and Adeniyi, 2024).

Although prior research has examined Generative AI and knowledge-mapping tools separately, there is still limited empirical evidence on how these approaches operate within the same learning environment. In particular, existing studies do not sufficiently explain how their different learning mechanisms influence student performance and conceptual understanding when applied to the same instructional tasks. Generative AI tools offer ease of use, immediate feedback, and personalized learning experiences, making them suitable for introductory learning and quick content generation. However, their potential for inaccuracy and over-reliance must be carefully managed. Mytelemap, on the other hand, promotes deep conceptual understanding and critical thinking through active knowledge organization, though it requires a greater investment of time and effort from learners. Based on recent experimental findings, educators may consider using both tools in tandem to leverage their complementary strengths and provide a balanced, effective learning experience.

Taken together, the literature suggests that Generative AI and knowledge-mapping tools support learning through fundamentally different mechanisms. Generative AI primarily facilitates rapid access to information, adaptive feedback, and task-oriented assistance, which can reduce cognitive load during early learning stages. In contrast, knowledge-mapping tools emphasize active knowledge construction, conceptual organization, and metacognitive engagement. However, existing studies have largely examined these approaches independently, focusing either on AI-supported learning or visualization-based tools. As a result, there remains limited empirical understanding of how these different forms of learning support function within the same instructional context, particularly in programming education where both procedural skills and conceptual understanding are critical.

This gap motivates the present study, which examines how Generative AI and knowledge-mapping approaches support programming learning when used within the same course context. By focusing on both performance and learner perceptions, the study aims to provide a more integrated understanding of how these complementary learning supports function in practice.

3. Methodology

3.1 Data Collection

The experiment was designed to examine differences in how Generative AI and Mytelemap support student learning in helping students learn through knowledge mapping. The study involved 30 first-year undergraduate students from the "Object Oriented Programming" course at Prince of Songkhla University International College, Thailand. An a priori power analysis was conducted using GPower 3.1 (Buchner, Faul and Erdfelder, 2024) to ensure an adequate sample size of N = 30, which was deemed sufficient for detecting large effect sizes with 95% power at the 5% level of significance. The sample size (N = 30) and the focus on a single course context should be considered when interpreting the findings. These factors may limit the generalizability of the results to other disciplines, institutions, or learner populations. The study was granted ethical approval by the University's Ethics Institutional Review Board under reference PSU IRB 2024-LL-Uic-019 (Internal).

The overview of experimental design is shown in Figure 2. The study took place over the final four weeks of the course, with students participating in one hour-long session per week. Each week, the students sketched their own knowledge mapping and used Generative AI or Mytelemap to learn and complete the assigned tasks. The learners were divided into two cohorts: Cohort A used Generative AI, while Cohort B used Mytelemap. The students were assigned specific tasks to complete using the respective tool, with one of the tasks being to complete "C# Methods" assignments. Mytelemap is a research-based knowledge-mapping environment developed to support structured conceptual learning through visualization. While its deployment is currently limited to the study context, it represents a class of knowledge-mapping tools designed to facilitate active knowledge construction. The purpose of including Mytelemap in this study is not to evaluate its general adoption, but to examine how this type of learning support functions in comparison with AI-based approaches within the same instructional setting.

In addition to using the tools, the students were also asked to complete surveys at the end of each session. These surveys gathered their opinions on the use of knowledge mapping in both tools, focusing on factors such as effectiveness, engagement, ease of use, learning enhancement, motivation, overall satisfaction, relevance, clarity, organization, comprehensiveness, usability, and whether the tools were up to date. The surveys helped measure the overall effectiveness of Generative AI and Mytelemap in supporting the learning process.

Furthermore, participants were assigned to complete learning tasks each week after using both tools. These tasks were designed to evaluate their understanding and application of the knowledge mapping tools. The students' task performance was assessed and compared between the two tools to determine their impact on learning outcomes. Task scores were recorded for each cohort to assess how effectively each tool supported the students' learning.

At the end of the study, the learners were provided with a URL to access the study findings and see how their data was used in the research. The study aimed to gather both quantitative data from task scores and qualitative data from student surveys, offering a comprehensive evaluation of the tools' effectiveness in enhancing learning outcomes. Through this design, the experiment compared both the performance and satisfaction of students using Generative AI and Mytelemap.

Experimental Design			Doc7
	E1 **Generative AI**	**E2** **Mytelemap**	**Duration (mins)**
Pre-activities			
Session 0. Introduction to the experiment.	N = 30 Purpose, duration, etc.		5
Introduction to the ideas.	N = 30 Mapping and getting study links activites		15
Explain tools.	N = 30 MyTeleMap & Generative AI		15
Treatment activities			
Session 1. Undertake task: "**C# Methods**".	N = 15 Cohort A	N = 15 Cohort B	60
Measure opinion	N = 30 Resulting task scores		10
Session 2. Undertake task: "**C# Arrays**".	N = 15 Cohort B	N = 15 Cohort A	60
Measure opinion	N = 30 Resulting task scores		10
Session 3. Undertake task: "**C# Lists**".	N = 15 Cohort B	N = 15 Cohort A	60
Measure opinion	N = 30 Resulting task scores		10
Session 4. Undertake task: "**C# Strings**".	N = 15 Cohort A	N = 15 Cohort B	60
Measure opinion	N = 30 Resulting task scores		10
Post-activities			
Session 5. Debrief	N = 30 Opinions on Tools Thanks, etc.		5
Options			
Score paper.			
Score products.			

Figure 2: Experimental Design: Overview of the Four-week Alternating Tool Use and Survey Timeline

After using the assigned tool each week, participants were asked to evaluate their satisfaction based on a series of statements related to their experience with the tools. The statements for Mytelemap and Generative AI were as follows, where the tool name was inserted at the position indicated:

- Eu: [Tool name] was easy to use for generating knowledge mappings.
- Ev: [Tool name] effectively helped in visualizing knowledge mappings.
- Ar: The knowledge mappings generated by [Tool name] were accurate and reliable.
- Al: The study materials links suggested by [Tool name] aligned well with my learning objectives.
- Ud: The study materials suggested by [Tool name] were current and up to date.
- Cc: [Tool name] allowed me to customize the criteria for suggesting study materials links.

The statements for Generative AI were:

- Ahs: Generative AI provided high quality of study contents.
- Ars: The study contents suggested by Generative AI were highly relevant to my study topics.
- Ats: I trusted the study contents suggested by Generative AI.

Participants rated each of the above statements on a Likert scale, from 1 (Strongly Disagree) to 5 (Strongly Agree). This evaluation assessed the learners' overall satisfaction with both tools in various aspects such as ease of use, visualization effectiveness, alignment with learning objectives, and content relevance. Although no single pre-existing validated instrument fully captured the constructs relevant to both Generative AI usage and knowledge-mapping tools, the questionnaire was adapted from scales previously used in technology acceptance and learning environment studies (e.g., usefulness, clarity, relevance, and satisfaction). At the end of the study some limited evidence was gathered to examine the reliability of the questionnaire. Participants encountered the questionnaire four times, twice in the context of using Generative AI, and twice in the context of using Mytelemap. These encounters were treated as data for calculating Cronbach's Alpha by analogy with the calculation of test-retest reliability. The resulting Cronbach's Alpha value (α = 0.23) indicates low internal consistency according to conventional methodological standards. This suggests that the questionnaire items do not measure a single coherent construct. This result should be interpreted in light of the study design, as the "retest" second encounter involved a different programming task, and the questionnaire was intentionally designed to capture multiple dimensions of learner perception (e.g., usability, engagement, and content relevance) rather than a single underlying construct. As such, Cronbach's Alpha may not be an appropriate measure of reliability for this instrument. Accordingly, the questionnaire results are interpreted as descriptive indicators of different aspects of learner experience rather than as a unified scale.

The questionnaires, class exercises and mapping scores criteria were reviewed by 2 independent experts to confirm face validity, clarity, and content relevance and were found to be clear and relevant measures. The experts included one from an external organization who had a research background similar to the investigator, and another was the head of curriculum, who was responsible for curriculum development and management. The experts used a 5-point Likert scale (1 = Not at all, 2 = Slightly, 3 = Moderately, 4 = Very, and 5 = Completely) in making 117 ratings, summarized in Table 1. All ratings were 4, 4.5, or 5; 89% of Expert 1 ratings were "5", 83% for Expert 2. All items rated less than 5 were revised before deployment.

Table 1: Expert 1 × Expert 2 Crosstabulation of 117 Ratings of Face Validity, Clarity, And Content Relevance

		Expert 2 ratings			Total
		4	4.5	5	
Expert 1 ratings	4	0	0	2	2
	4.5	0	0	11	11
	5	14	6	84	104
Total		14	6	97	117

Chi-squared (4) = 3.02, p = .55 for the crosstabulation. Some of the expected frequencies were less than 5, but because the Chi-square was not significant, no correction for continuity was made. This finding suggests that Expert 1's pattern of ratings was not significantly different from that of Expert 2.

This expert review procedure aligns with common validation practices in exploratory educational technology research, providing justification for using a customized questionnaire tailored to the unique features of the two tools.The learning tasks were based on weekly topics related to key programming concepts covered in the course, including programming classes, objects, encapsulation, and inheritance. In completing these tasks, students applied their knowledge of programming and object-oriented design.

The learning tasks were assessed using a set of criteria to evaluate the quality and effectiveness of the work produced by the participants:

- Rrw (Relevance to Real World): This criterion assessed how well the chosen scenario or problem related to the participants' daily life or interests. The criterion mark was based on how applicable and realistic the problem was in the context of real-world situations.
- Cor (Correctness): This evaluated whether inheritance was correctly implemented and functioned as intended. The participants' code was examined and marked according to the proper application of core programming concepts, especially inheritance.
- Cqu (Code Quality): This criterion focused on whether the code was clean, well-organized, and well-documented. The mark was based on clear and maintainable code that followed good programming practices and included adequate comments to explain their design choices.
- Ref (Reflection): The participants were assessed on how well they articulated their design choices and learning experiences. The mark was based on their explanation of the decisions they made during the task and reflection on how the concepts learned were applied in their solution.
- Tes (Testing): on the criterion assessed the thoroughness of the test cases provided. The mark was based on adequate and comprehensive testing of the correctness of the implementation.

This approach ensures that each participant's task was evaluated comprehensively, considering both the technical aspects of programming and the depth of their understanding and reflection on the learning process. A total mark was calculated as Tot (Total mark), weighted 20% Rrw, 30% Cor, 20% Cqu, 10% Ref, and 20% Tes.

Assessment and marking was carried out by an investigator blind to the identity of the participant and their experimental treatment, and used a detailed scoring rubric to keep scoring as consistent as possible.

Students were assigned to Cohort A (Generative AI first) and Cohort B (Mytelemap first) based on class attendance order during the first week; no pre-existing academic criteria were used for assignment. The counterbalanced ABBA and BAAB design mitigated order effects by ensuring that every student used both tools and allowed a within-subjects analysis of their data. Pre-tests and post-tests were not administered in this short four-week intervention; instead, weekly learning tasks served as performance indicators. Pre-tests and post-tests were not administered in this study; instead, learning outcomes were inferred from weekly task performance. This limits the ability to draw strong causal conclusions regarding learning gains attributable to each tool.

As the study was conducted within a natural classroom setting and did not include full experimental controls, the findings should be interpreted as indicative of relationships rather than definitive causal effects.

3.2 Data Analysis

As illustrated in Figure 2, each student used Mytelemap and Generative AI twice, termed their First Encounter in Week 1 or 2, and Second Encounter in week 3 or 4. Four sets of analyses of variance were conducted using IBM SPSS version 29. The first set examined the Total mark in a univariate repeated measures design; the second set investigated the five components of the Total mark (Rrw, Cor, Cqu, Ref, Tes) in a multivariate repeated measures design; the third set investigated the six opinion variables (Eu, Ev, Ar, Al, Ud, Cc) in a multivariate repeated measures design; the fourth set investigated the three Generative AI opinion variables (Ahs, Ars, Ats). Effect sizes for significant effects are shown as Cohen's d.

3.2.1 Total mark

Table 2 shows the interaction effect of Tool (Generative AI vs. Mytelemap) by Encounter (First or Second) was not significant, while the main effects of Tool and of Encounter were significant.

Table 2: Anova summary table for Total mark

Source	SSs	df	MS	F	p	d
Encounter	480.00	1	480.00	9.97	0.004	0.58
Error(Encounter)	1396.00	29	48.14			
Tool	433.20	1	433.20	17.77	<0.001	0.77
Error(Tool)	706.80	29	24.37			
Encounter * Tool	38.53	1	38.53	0.85	0.37	
Error(Encounter*Tool)	1321.47	29	45.57			

Where a factor comprises two levels, a significant main effect may be directly interpreted by inspection of the mean values of the variable concerned. The main effect of Tool showed that the Generative AI mean Total Mark (80.6) was significantly larger than that for Mytelemap (76.8), and the main effect of Encounter showed that the Second Encounter mean Total Mark (80.7) was significantly larger than that at First Encounter (76.7). These findings suggest that both tools contributed to improved learning with their second use, and that Generative AI contributed more to the participants' marks on the learning tasks than Mytelemap.

3.2.2 Components of total mark

The five components of the Total mark (Rrw, Cor, Cqu, Ref, and Tes) were investigated to explore if the significant differences between Tools, and between Encounters, on Total Mark were associated with any particular component of that mark. Table 3 provides the multivariate analysis of the five components, where it is seen that the overall interaction effect of Tool * Encounter was not significant, while the overall main effects of Tool and Encounter were significant.

Table 3: Multivariate anova summary table for Rrw, Cor, Cqu, Ref, and Tes

Source	F	Hypothesis df	Error df	p
Encounter	6.24c	5	25	<0.001
Tool	3.84c	5	25	0.01
Encounter * Tool	1.60	5	25	0.20

Following the multivariate tests, univariate analyses were conducted on the main effects (only) of Tool and Encounter for the five components, as summarized in Table 4. It may be seen that the Tool main effect was significant for Cqu and Ref, and not significant for Rrw, Cor, and Tes. The Encounter main effect was significant for Ref and Tes, and not significant for Rrw, Cor, and Cqu.

Table 4: Univariate anova summary table for Rrw, Cor, Cqu, Ref, and Tes

Source	Measure	SS	df	MS	F	p	d
Tool	Rrw	0.68	1	0.68	3.51	0.07	
	Cor	0.41	1	0.41	2.22	0.15	
	Cqu	2.13	1	2.13	6.27	0.02	0.46
	Ref	3.68	1	3.68	5.59	0.03	0.43
	Tes	1.01	1	1.01	3.00	0.09	
Error(Tool)	Rrw	5.57	29	0.19			
	Cor	5.34	29	0.18			
	Cqu	9.87	29	0.34			
	Ref	19.07	29	0.66			
	Tes	9.74	29	0.34			
Encounter	Rrw	1.01	1	1.01	2.39	0.13	
	Cor	0.21	1	0.21	0.45	0.51	
	Cqu	0.13	1	0.13	0.28	0.60	
	Ref	2.41	1	2.41	6.75	0.01	0.47
	Tes	7.01	1	7.01	20.86	<.001	0.83
Error(Encounter)	Rrw	12.24	29	0.42			
	Cor	13.54	29	0.47			
	Cqu	13.87	29	0.48			
	Ref	10.34	29	0.36			
	Tes	9.74	29	0.34			

The main effect of Tool on Cqu showed that the Generative AI mean (3.98) was significantly larger than for Mytelemap (3.72), and on Ref showed that the Generative AI mean (4.02) was significantly larger than for Mytelemap (3.67). This suggests that the larger mean Total Mark for Generative AI than for Mytelemap was attributable to better code quality (Cqu) and to better reflection on the design choices (Ref), and that there was no significant difference between the tools on real world relevance (Rrw), correctness (Cor), or testing (Tes).

The main effect of Encounter on Ref showed that the Second Encounter mean (3.98) was significantly larger than that at First Encounter (3.70), and on Tes showed that the Second Encounter mean (3.95) was significantly larger than that at First (3.47). This suggests that the larger mean Total Mark at Second Encounter was attributable to better reflection on the design choices (Ref) and to better testing (Tes), and that there was no significant gain at Second Encounter attributable to code quality (Cqu), real world relevance (Rrw), or correctness (Cor).

3.2.3 *Opinions on Tools*

The six opinion variables (Eu, Ev, Ar, Al, Ud, Cc) were investigated to explore if they showed significant differences between Tools or Encounters. Table 5 provides the multivariate analysis of the six opinions, where it is seen that the overall interaction effect of Tool * Encounter was significant.

Table 5: Multivariate anova summary table for Eu, Ev, Ar, Al, Ud, Cc

Source	F	Hypothesis df	Error df	p
Encounter	1.50c	6	24	0.22
Tool	42.83c	6	24	<0.001
Encounter * Tool	5.40c	6	24	0.001

Univariate analyses were conducted on the six opinion variables, as summarized in Table 6. It may be seen that the interaction effect of Tool * Encounter was significant for Ev and Ud, hence their main effects are not interpreted for these variables. The Tool effect was significant for Eu, Ar, and Al, and not for Cc. The Encounter effect was significant for Eu, and not for Ar, Al, or Cc.

Table 6: Univariate anova summary table for Eu, Ev, Ar, Al, Ud, Cc

Source	Opinion	SS	df	MS	F	p	d
Encounter	Eu	1.20	1	1.20	6.00	0.02	0.45
	Ev	0.08	1	0.08	0.24	0.63	
	Ar	0.07	1	0.07	0.38	0.54	
	Al	0.03	1	0.03	0.16	0.69	
	Ud	0.30	1	0.30	0.54	0.47	
	Cc	0.30	1	0.30	0.57	0.46	
Error(Encounter)	Eu	5.80	29	0.20			
	Ev	9.17	29	0.32			
	Ar	5.67	29	0.20			
	Al	5.97	29	0.21			
	Ud	16.20	29	0.56			
	Cc	15.20	29	0.52			
Tool	Eu	22.53	1	22.53	56.99	<0.001	1.38
	Ev	46.87	1	46.87	109.85	<0.001	1.91
	Ar	35.21	1	35.21	156.08	<0.001	2.28
	Al	2.70	1	2.70	4.53	0.04	0.39
	Ud	0.83	1	0.83	1.77	0.19	
	Cc	0.03	1	0.03	0.08	0.77	
Error(Tool)	Eu	11.47	29	0.40			
	Ev	12.38	29	0.43			
	Ar	6.54	29	0.23			
	Al	17.30	29	0.60			
	Ud	13.67	29	0.47			
	Cc	11.47	29	0.40			
Encounter * Tool	Eu	0.53	1	0.53	2.07	0.16	
	Ev	3.01	1	3.01	16.64	<0.001	0.74
	Ar	0.07	1	0.07	0.59	0.45	
	Al	1.20	1	1.20	3.22	0.08	
	Ud	4.03	1	4.03	12.36	0.001	0.64

Source	Opinion	SS	df	MS	F	p	d
	Cc	0.30	1	0.30	0.61	0.44	
Error(Encounter*Tool)	Eu	7.47	29	0.26			
	Ev	5.24	29	0.18			
	Ar	3.67	29	0.13			
	Al	10.80	29	0.37			
	Ud	9.47	29	0.33			
	Cc	14.20	29	0.49			

A significant interaction effect requires the investigation of the simple main effects of a factor at each level of the other factor. Where each factor comprises two levels, simple main effects may be directly tested by pairwise comparisons. Table 7 provides the pairwise comparisons between First and Second Encounters for mean Ev and Ud opinions of participants when they used Generative AI and when they used Mytelemap, and Table 8 provides the pairwise comparisons between Generative AI and Mytelemap for mean Ev and Ud opinions of participants at First and Second Encounters.

While it can be seen that the Second Encounter using Generative AI was rated significantly higher on help with visualizing knowledge mappings than First Encounter, and Second Encounter was rated significantly lower than First Encounter using Mytelemap, it is clear that visualizing knowledge mappings using Mytelemap was rated significantly higher than Generative AI at both First and Second Encounter.

It can be seen that the Second Encounter using Generative AI was rated not significantly different on up-to-date study materials compared with First Encounter, while Second Encounter was rated significantly higher than First Encounter on up-to-date materials using Mytelemap. Study materials were rated as more up to date on First Encounter when using Generative AI than Mytelemap, and the difference between the tools was rated as not significant at Second Encounter.

Table 7: Pairwise comparisons between First and Second Encounters for mean Ev and Ud

Opinion	Tool	First Encounter	Second Encounter	Mean Diff	p	d
Ev	Generative AI	3.23	3.60	-0.37	0.01	0.65
	Mytelemap	4.80	4.53	0.27	0.04	0.48
Ud	Generative AI	4.60	4.33	0.27	0.07	0.36
	Mytelemap	4.07	4.53	-0.47	0.02	0.63

Table 8: Pairwise comparisons between Generative AI and Mytelemap for mean Ev and Ud

Opinion	Encounter	Generative AI	Mytelemap	Mean Difference	p	d
Ev	First	3.23	4.80	-1.57	<0.001	2.78
	Second	3.60	4.53	-0.93	<0.001	1.64
Ud	First	4.60	4.07	0.53	0.01	0.71
	Second	4.33	4.53	-0.20	0.18	

The main effect of Encounter showed a significantly higher mean rating on Second Encounter of ease of use (4.13) than on First Encounter (3.93). There was no significant difference between First and Second Encounter mean opinion ratings for accuracy and reliability (Ar), alignment with learning objectives (Al), or customizable criteria (Cc) (overall means 4.11, 4.40, 4.42 respectively).

The main effect of Tool showed a significantly higher mean rating of ease of use (Eu) for Mytelemap (4.47) than for Generative AI (3.60) and a significantly higher rating for accuracy and reliability (Ar) (4.65 vs 3.57), but a significantly lower rating for alignment with learning objectives (Al) (4.25 vs 4.55). There was no significant difference between Generative AI and Mytelemap mean opinion ratings for customizable criteria (Cc) (overall mean 4.42).

The mean opinion ratings all showed positive agreement, and were significantly higher than a "Neither agree nor disagree" rating. The results suggest that, in general, Mytelemap provided significantly higher ratings than

Generative AI, if not at First Encounter then at the Second Encounter, on visualizing knowledge mappings, ease of use, accuracy and reliability, and up-to-date materials. The exceptions were that Generative AI rated more highly for alignment with learning objectives, and that both tools were rated as equally customizable.

3.2.4 Opinions on Generative AI

The three Generative AI opinion variables (Ahs, Ars, Ats) were investigated to explore if they showed significant differences between Encounters. Table 9 provides the multivariate analysis of the opinions, where it is seen that the Encounter effect was significant.

Table 9: Multivariate anova summary table for Ahs, Ars, Ats

Source	F	Hypothesis df	Error df	p
Encounter	3.06	3	27	0.05

Univariate analyses were conducted on the three opinion variables, as summarized in Table 10. It may be seen that the Encounter effect was significant for Ats, but not for Ahs or Ars.

Table 10: Univariate anova summary table for Ahs, Ars, Ats

Source	Opinion	SS	df	MS	F	p	d
Encounter	Ahs	0.02	1	0.02	0.07	0.79	
	Ars	0.27	1	0.27	1.00	0.33	
	Ats	1.35	1	1.35	7.60	0.01	0.50
Error(Encounter)	Ahs	6.48	29	0.22			
	Ars	7.73	29	0.27			
	Ats	5.15	29	0.18			

The mean opinion ratings all showed positive agreement, and were significantly higher than a "Neither agree nor disagree" rating. Mean trust in the study contents (Ats) of the Generative AI tool significantly increased from First Encounter (4.3) to Second (4.6), while there was no significant change in mean opinion on quality (Ahs) or relevance (Ars) of the study contents of the Generative AI tool (overall means 4.7 and 4.7 respectively).

3.2.5 Grade Point Average (GPA)

The relationship between GPA and Total Mark was explored by correlating the participants' GPA with their mean Total Mark for Generative AI and Mytelemap, and with the difference in these two means. As would be expected, GPA correlated highly with both Generative AI and Mytelemap Total Marks ($r(28) = .66$ and $.70$, both $p < .001$), but more interestingly did not correlate with the difference in mark, $r(28) = 0.04$, $p = .83$, This suggests that whether a participant found one or other tool more useful was not particularly related to their GPA.

4. Results and Discussion

These results align strongly with theoretical expectations from Cognitive Load Theory and Constructivist pedagogy. Generative AI appears to reduce extraneous cognitive load by providing immediate clarification and simplified explanations, which helps novices overcome early programming barriers. This may help explain patterns observed in task performance and ease-of-use ratings. Conversely, Mytelemap requires learners to actively construct conceptual relationships, consistent with constructivist principles of meaning-making and schema integration. The cognitive effort required to build, revise, and navigate mappings appears to foster deeper reflection and conceptual clarity, particularly evident in improvements in the Reflection and Code Quality components during later encounters. This reinforces the pedagogical value of knowledge visualization when learning goals involve abstraction, synthesis, and higher-order understanding.

These findings are consistent with the theoretical perspectives discussed earlier, particularly in relation to Cognitive Load Theory and constructivist learning principles. Generative AI appears to support learning by providing immediate clarification and reducing cognitive effort during early stages, while Mytelemap encourages active knowledge construction and conceptual organization through visual mapping.

The results indicate differences in how students engaged with the two learning environments. Student performance was higher in later encounters with both tools, suggesting increased familiarity over time. However, differences emerged in usability, perceived reliability, and support for conceptual understanding.

Generative AI was associated with higher ratings for ease of use, responsiveness, and alignment with immediate learning needs. Students valued the ability to obtain quick explanations and examples, particularly when encountering new programming concepts. This contributed to higher engagement in early sessions. However, some limitations were observed in terms of content accuracy, particularly in more complex topics, where learners reported inconsistencies and the need to verify information using additional sources.

In contrast, Mytelemap appeared to support deeper conceptual engagement, particularly during later stages of the study. The tool encouraged learners to structure relationships between concepts, which was reflected in stronger performance in tasks involving reflection and code quality. Students also reported increased confidence in the relevance and organization of learning materials over time. However, initial usability challenges were noted, as some learners required time to become familiar with the interface and workflow.

The findings suggest that the two tools support different aspects of learning. Generative AI appears to facilitate rapid task completion and immediate clarification, while Mytelemap encourages more structured knowledge organization and reflection. This indicates that the tools may play complementary roles in programming learning. The two approaches differ in their underlying design and knowledge representation; therefore, the findings should be interpreted in terms of their complementary roles rather than direct equivalence.

While the general distinction between rapid AI-supported assistance and deeper conceptual learning through visualization may appear intuitive, the contribution of this study lies in providing empirical evidence of how these differences manifest within the same instructional context. Rather than simply confirming an assumed difference, the findings show how these learning mechanisms operate under comparable conditions, using both performance data and learner perceptions. In particular, the results indicate that the benefits of each approach vary across stages of learning and types of tasks, offering a more nuanced understanding of how different forms of learning support function in practice. This is particularly relevant given the increasing use of both AI-driven tools and visualization-based approaches in programming education, where understanding how different forms of support influence learning processes remains an open research question.

Overall, these results should be interpreted within the context of the study design. The findings contribute to the e-learning field by providing empirical insight into how different forms of learning support function within the same instructional environment. By highlighting the complementary and context-dependent roles of Generative AI and knowledge mapping, the study offers a more nuanced perspective on integrating adaptive and structured learning approaches in programming education.

5. Conclusion

This study examined how Generative AI and knowledge mapping approaches are associated with differences in student learning within an object-oriented programming course. The findings reveal that while Generative AI excels in providing quick, personalized learning support, it is limited by concerns over content accuracy. Mytelemap, as a knowledge-mapping approach, appears to support deeper conceptual understanding through structured visualization and active knowledge organization, although it presents usability challenges for new users. Framing these findings through Cognitive Load Theory and Constructivist principles suggests that Generative AI is effective for reducing extraneous load and facilitating quick comprehension, whereas Mytelemap supports deeper schema development and reflective learning. This theoretical integration highlights the potential of hybrid instructional approaches that intentionally sequence AI-mediated scaffolding with visualization-based consolidation activities. As the two approaches differ in their underlying pedagogical design and knowledge representation, they are not treated as directly equivalent in this study. Instead, the comparison focuses on how different forms of learning support—Generative AI and knowledge mapping—are associated with variations in student learning processes and experiences within the same instructional context.

Rather than viewing one approach as superior, the findings suggest that each supports different aspects of learning and may be more effective under different conditions. Generative AI appears to support rapid clarification and early-stage learning tasks, while knowledge-mapping approaches appear to support conceptual organization and reflective understanding. This distinction provides a basis for designing instructional strategies that align specific tools with particular learning objectives and stages. These findings provide a foundation for future research by identifying how different forms of learning support function within a shared instructional environment. This enables further investigation into how hybrid learning designs can be structured, sequenced, and evaluated across different contexts and disciplines. However, this study has several limitations. First, student perceptions were measured using self-report surveys, which may be subject to bias. Second, participants were more familiar with conversational interfaces than with knowledge mapping tools, potentially affecting initial usability ratings. Third, the intervention spanned only four weeks, limiting the ability to observe long-term learning gains. Finally, although students experienced both tools, the study design did not include full pre/post testing, which constrains causal interpretations regarding learning improvement. Future research should incorporate longer interventions and triangulated performance measures. Future research should explore other domains and more diverse learner populations. In addition, it should investigate the integration of Generative AI into knowledge mapping platforms like Mytelemap, potentially allowing AI-generated suggestions to enhance map creation. Additionally, expanding the study across diverse disciplines and student populations would help validate these findings and explore broader applications. A hybrid learning model that leverages both tools' strengths may offer a promising path forward in technology-enhanced education.

Acknowledgement

This work was supported by Prince of Songkla University (Ref. No. UIC6901116S), Thailand Science Research and Innovation (TSRI) [Grant numbers FRB690047/0230], and National Science, Research and Innovation Fund (NSRF).

Ethical Approval: This study was reviewed and approved by the Prince of Songkla University Ethics Institutional Review Board under reference number PSU IRB 2024-LL-Uic-019 (Internal).

Availability of Data and Materials: The datasets generated and analyzed during the current study are available from the corresponding author upon reasonable request.

Declarations: Generative AI was used in this study as a subject of investigation and not as a tool for data generation or analysis. AI-assisted tools were used only for minor language editing. All research design, analysis, and conclusions were developed by the authors, who take full responsibility for the content.

References

Bender, E. M., Gebru, T., McMillan-Major, A. & Shmitchell, S. On the dangers of stochastic parrots: Can language models be too big?? Proceedings of the 2021 ACM conference on fairness, accountability, and transparency, 2021. 610-623.

Buchner, A., Faul, F. & Erdfelder, E. 2024. *G*Power statistical power analysis program* [Online]. Available: http://www.psycho.uni-duesseldorf.de/aap/projects/gpower/ [Accessed 1/3/2026].

Eden, C. A., Chisom, O. N. & Adeniyi, I. S. 2024. Integrating AI in education: Opportunities, challenges, and ethical considerations. *Magna Scientia Advanced Research and Reviews,* 10**,** 006-013.

Hutson, J., Lively, J., Robertson, B., Cotroneo, P. & Lang, M. 2023. Expanding Horizons: AI tools and workflows in Art Practice. *Creative Convergence: The AI Renaissance in Art and Design.* Springer.

Khreisat, M. N., Khilani, D., Rusho, M. A., Karkkulainen, E. A., Tabuena, A. C. & Uberas, A. D. 2024. Ethical implications of AI integration in educational decision making: Systematic review. *Educational Administration: Theory and Practice,* 30**,** 8521-8527.

Nitchot, A. & Gilbert, L. 2024a. Expressing learners' competences through the use of pedagogically-informed knowledge mapping: case study in constructing a holographic projector. *Journal of Research in Innovative Teaching & Learning*.

Nitchot, A. & Gilbert, L. 2024b. Supporting learning and motivation with knowledge mapping in programming courses. *Education and Information Technologies***,** 1-24.

Nitchot, A. & Gilbert, L. 2025. Comparison of Mytelemap and MindMeister in using competence maps for self-learning. *Technology, Pedagogy and Education,* 34**,** 69-89.

Nitchot, A., Wettayaprasit, W. & Gilbert, L. 2019. Assistive tool for constructing knowledge structures and suggesting related study materials links. *Education and Information Technologies,* 24**,** 219-230.

Rahayu, N. W., Ferdiana, R. & Kusumawardani, S. S. 2022. A systematic review of ontology use in E-Learning recommender system. *Computers and Education: Artificial Intelligence,* 3**,** 100047.

Saqr, R. R., Al-Somali, S. A. & Sarhan, M. Y. 2024. Exploring the acceptance and user satisfaction of AI-driven e-learning platforms (Blackboard, Moodle, Edmodo, Coursera and edX): an integrated technology model. *Sustainability,* 16**,** 204.

Singh, A., Lakhera, G., Ojha, M., kumar Mishra, A. & Nain, A. 2025. Balancing Innovation With Responsibility: Ethical Dimensions of AI in Revolutionizing E-Learning. *Ethical Dimensions of AI Development.* IGI Global.

Strielkowski, W., Grebennikova, V., Lisovskiy, A., Rakhimova, G. & Vasileva, T. 2024. AI-driven adaptive learning for sustainable educational transformation. *Sustainable Development*.

Sweller, J. 1994. Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction,* 4**,** 295-312.

Ward, B., Bhati, D., Neha, F. & Guercio, A. Analyzing the impact of AI tools on student study habits and academic performance. 2025 IEEE 15th Annual Computing and Communication Workshop and Conference (CCWC), 2025. IEEE, 00434-00440.

Were, M. C. 2022. Challenges in digital medicine applications in under-resourced settings. *Nat Commun,* 2022**,** 3020.

Xodabande, I., Atai, M. R. & Hashemi, M. R. 2024. Exploring the effectiveness of mobile assisted learning with digital flashcards in enhancing long-term retention of technical vocabulary among university students. *Journal of Computers in Education***,** 1-26.

Apriori-based Analysis of Learned Helplessness in Mathematics Tutoring: Behavioral Patterns by Level, Intervention, and Outcome

John Paul P. Miranda
Pampanga State University, Philippines

jppmiranda@pampangastateu.edu.ph

https://doi.org/10.34190/ejel.24.2.4602

Abstract: This study applied the Apriori algorithm to analyze behavioral interaction patterns associated with learned helplessness (LH) in mathematics tutoring system logs. Interaction data were examined across three dimensions: LH level (low vs. high), system-based intervention (with vs. without), and problem-solving outcomes (solved vs. unsolved). The analysis of the complete dataset showed that skipping problems without using hints was the most frequent pattern linked to unsolved outcomes, while persistence behaviors such as not skipping were less dominant overall. Comparisons by LH level showed that low-LH students had stronger links between problem solving and not skipping, as well as positive associations between hint use and solved outcomes. High-LH students showed more avoidance patterns, with skipping strongly tied to unsolved outcomes. In the comparison of system-based intervention conditions, students without intervention had the highest lift for persistence–success links, while the with-intervention group had stronger patterns involving skipping behaviors leading to unsolved outcomes. Outcome-specific analysis showed that not skipping was consistently associated with solved problems across all groups, while skipping without hints predicted unsolved outcomes. Practical implications and recommendations are discussed.

Keywords: Student engagement, Digital learning, Problem-solving strategies, Learning analytics, Help-seeking behavior, Educational data mining

1. Introduction

Students' behavior while working on mathematics problems reflects their motivation, persistence, and approach to challenges (Albay, 2020). In some cases, these behaviors indicate learned helplessness (LH), a state in which learners expect failure and reduce their effort (Amadi, Agi, & Nwoke, 2020; Hwang, 2019; Yates, 2009). In mathematics, LH can lead to giving up after mistakes, avoiding difficult problems, or skipping opportunities to seek help (Biber and Biber, 2014; Gürefe and Bakalım, 2018). Intelligent tutoring systems (ITS) have been developed to guide students through practice tasks, track their actions, and offer targeted support. These systems have been shown to help learners by adapting to their needs (Muangprathub, Boonjing, & Chamnongthai, 2020; Spitzer and Moeller, 2023), but they can also reveal patterns of negative behaviors that limit learning gains (Fancsali, 2014; Namukasa et al., 2023; du Plooy, Casteleijn, & Franzsen, 2024; Yang et al., 2022).

The behavioral patterns associated with LH in tutoring systems connect to several motivational frameworks. Attribution theory holds that students who attribute failure to fixed, uncontrollable causes tend to disengage rather than persist (Hwang, 2019; Weiner, 1986). In a tutoring context, this appears as skipping problems after mistakes rather than reattempting them. Self-determination theory adds that when students' need for competence goes unmet, their motivation to engage weakens (Ryan and Deci, 2000). Students who avoid hints may do so partly because prior help-seeking produced no felt sense of progress. Self-regulated learning research treats help-seeking as a deliberate strategy used by effective learners, and links help avoidance to weaker self-regulation and lower achievement (Yang, 2023; Zimmerman, 2000). Skipping and non-use of hints in tutoring logs are therefore not isolated actions because they carry motivational meaning that LH theory can help interpret.

Research on LH has often relied on surveys or experimental tasks, with fewer studies using detailed interaction logs from tutoring systems (Miranda et al., 2025; Yates, 2009). While data mining methods can uncover patterns in large datasets, the Apriori algorithm has frequently been applied to examine how combinations of student behaviors (Bringula et al., 2025; Fu, Ren, & Lin, 2025; Tang et al., 2024) but not in the context of LH particularly in mathematics. There is also limited work comparing these patterns between learners with different LH levels or between those with and without system-based interventions. Without this information, it is difficult to design support features that respond to the specific ways different students engage or disengage.

Cite this article: Miranda, J.P.P. 2026. "Apriori-Based Analysis of Learned Helplessness in Mathematics Tutoring: Behavioral Patterns by Level, Intervention, and Outcome", *Electronic Journal of e-Learning*, 24(2), pp 148-156, https://doi.org/10.34190/ejel.24.2.4602

This study applies the Apriori algorithm to interaction logs from a mathematics tutoring system to identify patterns in student behaviors. Specifically, it aims to: (1) examine the distribution of behavioral indicators such as mistakes, hint use, skipping, and solution status; (2) identify frequent behavioral patterns in the complete dataset; (3) compare these patterns between students with low and high LH levels; (4) compare patterns between students with and without system-based intervention; and (5) determine which patterns are associated with solved and unsolved problems in each group. The results are expected to provide practical insights for improving adaptive features in tutoring systems so they can better support persistence and reduce avoidance behaviors in mathematics learning.

2. Methodology

2.1 Dataset

The dataset for this study came from AES (Adaptive Equation Sensei) (Fig. 1), a mathematics tutoring system used by Grade 8 students in the Philippines (Miranda and Bringula, 2023; Miranda et al., 2025) based from (Bringula et al., 2015). Data originally came from two separate groups: students who used the system without intervention and those who used it with system-based interventions. The system-based intervention consisted of automated hints triggered by the system when a student made an error, motivational messages displayed during the session, and prompts designed to encourage continued engagement with the problem. Students in the without-intervention group used the same AES system but did not receive these features. Both datasets were stored together after collection for unified analysis. The complete dataset contained 3,696 interaction sessions generated by 246 students, with 113 students in the without-intervention group and 133 in the with-intervention group (Fig. 2). Data were collected in multiple periods between 2024 and early 2025. The initial pool included 193 students from eight public schools in the without-intervention group and 192 students from six different public schools in the with-intervention group.

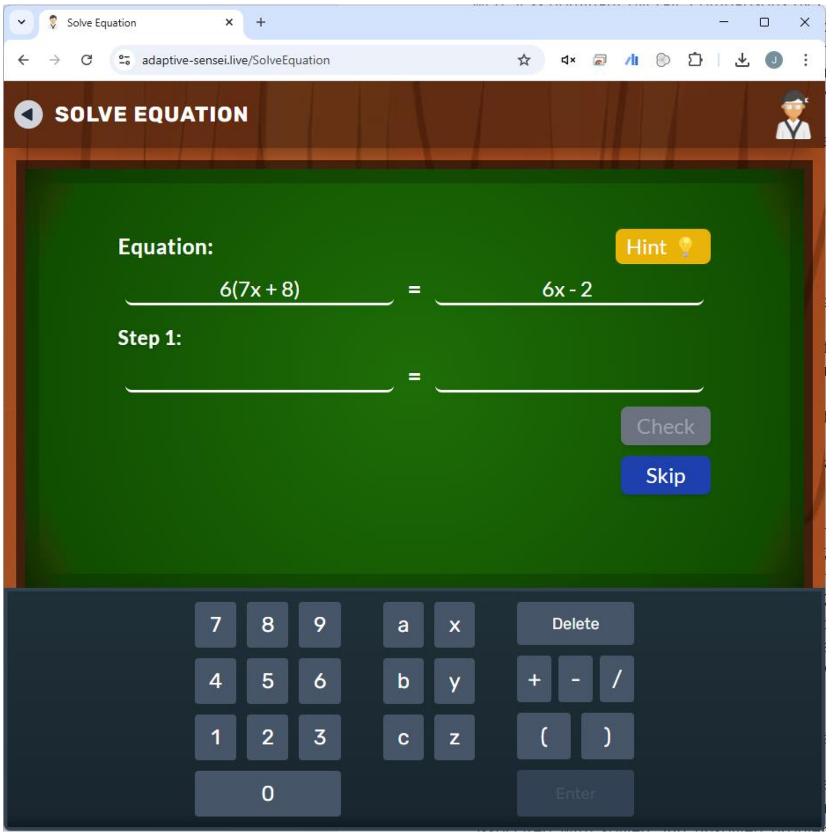


Figure 1: Adaptive Equation Sensei

	A	B	C	D	E	F	G	H	I	J	K
1	Account	MistakeO	HintUsed	Skipped	Status	TotalStep	TotalHints	TotalAnsw	TimeSpen	With Inte	Label
2	ABIS01	YES	NO	NO	UNSOLVED	2	0	11	466	YES	Low
3	ABIS01	NO	NO	NO	UNSOLVED	2	0	3	120	YES	Low
4	ABIS01	YES	NO	NO	SOLVED	1	0	3	117	YES	Low
5	ABIS01	YES	NO	YES	UNSOLVED	2	0	2	238	YES	Low
6	ABIS01	YES	NO	NO	UNSOLVED	1	0	3	179	YES	Low
7	ABIS02	YES	NO	YES	UNSOLVED	0	0	2	137	YES	High
8	ABIS02	YES	NO	NO	UNSOLVED	1	0	4	161	YES	High
9	ABIS02	YES	NO	NO	SOLVED	1	0	3	91	YES	High
10	ABIS02	NO	NO	NO	SOLVED	1	0	1	19	YES	High
11	ABIS02	NO	NO	NO	SOLVED	1	0	1	21	YES	High
12	ABIS02	YES	NO	YES	UNSOLVED	1	0	4	69	YES	High
13	ABIS02	YES	NO	YES	UNSOLVED	0	0	1	20	YES	High
14	ABIS02	NO	NO	NO	UNSOLVED	0	0	0	2	YES	High
15	ABIS02	NO	NO	NO	UNSOLVED	0	0	0	0	YES	High
16	ABIS02	YES	NO	NO	UNSOLVED	0	0	2	145	YES	High
17	ABIS03	YES	NO	YES	UNSOLVED	0	0	1	65	YES	High
18	ABIS03	NO	NO	YES	UNSOLVED	0	0	0	20	YES	High

Figure 2: Dataset exported to spreadsheet from the mobile tutoring system

2.2 Variables

The analysis used behavioral variables recorded by the tutoring system. For the purposes of this study, a session refers to a single continuous interaction with the AES system by one student, during which one or more problems were attempted, resulting in one row of logged behavioral data. These included binary indicators for mistake occurrence (*MistakeOccurred*), hint use (*HintUsed*), problem skipping (Skipped), and problem status (Status, solved or unsolved). Additional session-level counts included *TotalSteps*, *TotalHints*, and *TotalAnswerAttempts*. Two grouping variables were used: *With Intervention* (with vs. without) and *Label* (low vs. high LH). No demographic variables were included.

The LH label used to classify students into low and high groups was derived from a supervised machine learning classification model developed in prior work using the same AES system (Miranda and Bringula, 2023; Miranda and Bringula, 2025; Miranda et al., 2025). In those studies, LH was operationalized using Yates's (2009) 10-item teacher-rated scale, which served as the ground truth for classifying students into low and high LH groups based on observed behavioral indicators in mathematics. A Random Forest model was then trained on behavioral and academic features extracted from interaction logs, including problem-solving success rates, hint use frequency, and skipping behavior, and evaluated through 10-fold cross-validation, achieving 92% accuracy, a mean F1-score of 0.93, and a recall of 98% for high LH cases (Miranda and Bringula, 2025). In the present study, the binary output of this model was used to assign LH labels to students, allowing behavioral interaction patterns to be compared across low and high LH groups. This approach to operationalizing LH is consistent with the broader use of behavioral indicators as proxies for psychological constructs in educational data mining research, where self-report instruments are not always feasible within system-based data collection contexts (Fancsali, 2014; Yates, 2009).

2.3 Data Preparation and Analysis

The dataset was cleaned to remove blank or incomplete records. All retained variables were in binary or categorical form to fit the Apriori algorithm's requirements. Account identifiers were excluded from the pattern mining itself but retained for grouping by intervention status and LH level. Data were filtered to create subgroups for each analysis dimension: LH level, intervention status, and problem-solving outcome. The 3,696 sessions were contributed by 246 students, with an average of approximately 15 sessions per student. Because each session was treated as an independent transaction, repeated observations from the same students are present in the dataset, which has implications for how the results should be interpreted. Within-student dependencies are not accounted for, meaning that behavioral tendencies of individual students are reflected across multiple transactions rather than a single observation. As a result, students with more sessions contribute more heavily to the frequency counts that determine which *itemsets* meet the minimum support threshold. The patterns reported here therefore characterize session-level behavioral co-occurrences across the dataset rather than behavioral profiles of individual students, and findings should be interpreted accordingly rather than as statements about consistent tendencies within a given learner. This approach is consistent with session-level association rule mining in comparable studies (Bringula et al., 2025; Wang, Xiao, & Ma, 2022). Data cleaning and preparation were completed using Python in Jupyter Notebook, with supplementary inspection in spreadsheet software.

The analysis used the Apriori algorithm implemented through the *mlxtend.frequent_patterns* library in Python. The Apriori algorithm was selected over alternative data mining approaches for three reasons. First, it produces association rules expressed through support, confidence, and lift. These metrics are directly readable by educators, school administrators, and tutoring system designers who may not have a background in statistical modelling (Agrawal and Srikant, 1994; Wang, Xiao, & Ma, 2022). Predictive models such as logistic regression or decision trees produce probability estimates that require interpretation within a modelling framework. Apriori, by contrast, produces explicit if-then behavioral rules that can be communicated to practitioners and used directly to inform design decisions. Second, each session in the dataset is represented by a set of binary behavioral indicators. This session-level transactional structure is the format for which Apriori was designed. Third, the research objective is to identify which behavioral indicators co-occur within sessions linked to specific learner profiles and outcomes. It is not to reconstruct the order of individual actions within a session. Sequential pattern mining is suited for detecting ordered event chains in time-stamped logs (De et al., 2022; Real, Pimentel, & Braga, 2021; Zhang and Paquette, 2023). Because this study aggregates behavioral indicators at the session level, temporal ordering is not an analytical concern, and Apriori is the more appropriate method.

Data were first transformed into a transaction format using *TransactionEncoder* from the *mlxtend.preprocessing* module. Minimum thresholds for support, confidence, and lift were set to filter meaningful patterns (support ≥ 0.20, confidence ≥ 0.60, lift > 1). These thresholds were selected based on conventions established in educational data mining research and on the characteristics of the dataset. A minimum support of 0.20 means that a behavioral pattern must appear in at least 20% of sessions to be considered frequent. This threshold is appropriate for a dataset of 3,696 sessions because it ensures that reported patterns reflect the behavior of a meaningful proportion of the sample, rather than isolated cases (Hikmawati, Maulidevi, & Surendro, 2021; Papadogiannis, Wallace, & Karountzou, 2024). A minimum confidence of 0.60 was applied to retain only rules where the antecedent predicts the consequent in at least 60% of relevant transactions. This level reflects a moderately strong directional association and is consistent with thresholds used in comparable student behavior studies (Bringula et al., 2025; Fu, Ren, & Lin, 2025). A lift threshold greater than 1.0 ensures that the association between two behavioral indicators is stronger than what would be expected by chance alone, which is the minimum condition for a rule to carry practical meaning (Agrawal and Srikant, 1994; Sowan et al., 2025). Only the top 30 rules based on lift values were retained for detailed analysis. Association rules were generated for the entire dataset and separately for subgroups based on LH level, intervention status, and outcome. Descriptive statistics were computed using *pandas* in Python, while the generated *itemsets* and rules were saved to CSV files for review and interpretation.

To examine the stability of the reported patterns, a sensitivity check was conducted by re-running the Apriori algorithm across nine threshold combinations: minimum support values of 0.15, 0.20, and 0.25, crossed with minimum confidence values of 0.50, 0.60, and 0.70, while maintaining a consistent lift threshold greater than 1.0 across all subgroups. As shown in Table 1, the two primary avoidance-related rules, {Skipped} → {Unsolved} and {Skipped, HintUsed = No} → {Unsolved}, returned identical lift values across all nine combinations in every subgroup, indicating that these associations are not artifacts of the specific cutoff values selected for the primary analysis. The persistence-success rule, {Not Skipped} → {Solved}, was detected only in the without-intervention subgroup and only at a minimum confidence of 0.50, which is below the 0.60 threshold used in the primary analysis. The compound rule {Mistake, Skipped} → {Unsolved} was absent in the with-intervention group at a minimum support of 0.25 because its observed itemset support (0.231) fell below that cutoff.

Table 1: Lift values of key association rules by subgroup across all threshold combinations

Subgroup	{Skipped} → {Unsolved}	{Skipped, No Hint} → {Unsolved}	{Mistake, Skipped} → {Unsolved}	{Not Skipped} → {Solved}
Overall	1.244	1.244	1.244	-
Low LH	1.261	1.261	1.261	-
High LH	1.202	1.202	1.202	-
Without Intervention	1.252	1.252	1.252	2.847*
With Intervention	1.231	1.231	1.231**	-

Note. Lift values are identical across all nine threshold combinations (support ∈ {0.15, 0.20, 0.25}; confidence ∈ {0.50, 0.60, 0.70}; lift > 1.0). A dash indicates the rule did not meet the applicable threshold. * Emerged only at confidence = 0.50. ** Not detected at support = 0.25 (observed support = 0.231).

3. Results and Discussion

3.1 Distribution of Behavioral Indicators

The behavioral patterns differed between intervention and non-intervention groups, as well as between students with low and high levels of LH. Students in the intervention group were less likely to use hints (85.9%) than those in the non-intervention group (65.8%), while low LH students used hints more often (28.3%) than high LH students (21.1%). Mistake rates were comparable across groups, with high LH students showing a slightly higher frequency (44.4%) than low LH students (41.8%). Skipping behavior was more frequent in the intervention group (53.7%) compared to the non-intervention group (35.1%). This difference may reflect variation in problem-solving approaches rather than a direct effect of the intervention, as the groups consist of different students. Solve rates were generally low across all groups, with the non-intervention group showing a slightly higher proportion of solved problems (20.1%) than the intervention group (18.8%). Across LH levels, low LH students had a higher proportion of solved problems (20.7%) compared to high LH students (16.8%).

3.2 Behavioral Patterns Identified Using the Apriori Algorithm

Skipping without using hints was the most frequent pattern associated with unsolved problems (lift = 1.46). Mistakes followed by skipping also appeared frequently, suggesting that errors often led to disengagement. It should be noted that the Apriori algorithm identifies co-occurrence patterns within sessions, not the temporal order of events. Patterns described here as ‘mistakes followed by skipping’ reflect the co-presence of these behaviors within a session, not a verified sequence in which one event preceded the other. The direction implied by such descriptions is interpretive and should not be taken as evidence of a within-session behavioral chain.

It is also worth noting that skipping a problem does not necessarily indicate avoidance driven by helplessness. In some cases, students may skip because a problem exceeds their current working memory capacity, making it a response to cognitive overload rather than motivational withdrawal (Evans et al., 2024; Kuldas et al., 2014; Nuvvula, 2016; Sweller, 1988; Tsaparlis, 2021). Others may skip strategically to manage limited session time. This behavior is consistent with adaptive self-regulation rather than disengagement (Dang and Koedinger, 2020; Su et al., 2025). The AES system logs record whether a skip occurred but cannot distinguish why. The interpretation of skipping as avoidance in this study rests on its co-occurrence with unsolved outcomes and its association with high LH profiles, not on direct evidence of intent. Future work combining log data with think-aloud protocols or brief post-session surveys could help clarify the motivational basis of this behavior. Patterns linking not skipping with solved problems were present but less common, indicating that avoidance behaviors were more prominent than persistence.

Research supports the role of help-seeking and engagement in shaping learning outcomes. Adaptive help-seeking strategies improve performance, yet many online learners avoid using available assistance (Yang, 2023). Help avoidance, which includes deliberately not using hints, is negatively correlated with learning outcomes and is present in about 19% of students in ITS (Aleven et al., 2006; Li et al., 2024). Students who avoid help often show lower transfer learning scores (Li et al., 2024). Cognitive studies indicate that engaging in challenging activities, such as interleaved practice, enhances memory and problem solving (Samani and Pan, 2021). Motivation also plays a role; higher interest levels are linked to greater persistence and reduced effort avoidance (Song, Kim, & Bong, 2019). These suggest that the frequent skipping and non-use of hints in the dataset reflect effort avoidance and weak self-regulation, while persistence and help-seeking behaviors are associated with better problem-solving performance.

3.3 Differences Between Students with Low and High LH Levels

Low LH students showed strong persistence patterns, with solved problems linked to not skipping (lift = 2.33) and a positive link between hint use and solving problems (lift = 1.39). High LH students had stronger associations between skipping and unsolved problems (lift = 1.39), as well as skipping without mistakes and unsolved outcomes (lift = 1.37), reflecting greater avoidance tendencies. Studies on LH indicate that students attributing failure to uncontrollable causes are less persistent and less likely to seek help (Maier and Seligman, 2016). Help avoidance and ineffective strategies such as “wheel spinning” can reinforce feelings of helplessness (Beck and Gong, 2013; Li et al., 2024; Sideridis, 2003; Song, Kim, & Bong, 2019). In contrast, adaptive help-seeking is linked to improved performance and better coping with challenges (Li, Che Hassan, & Saharuddin, 2023). The patterns observed suggest that low LH students benefit from persistence and help use, while high LH students' avoidance behaviors correspond with unsolved outcomes.

3.4 Differences Between Students with and Without Intervention

Students without intervention exhibited the strongest persistence-success association, with solved problems linked to not skipping (lift = 2.85). They also showed a notable pattern connecting hint use to mistakes (lift = 1.77). In the intervention group, skipping was closely tied to unsolved problems (lift = 1.35), and mistakes in unsolved problems often preceded skipping (lift = 1.34). Research on ITS shows that feedback on help-seeking can encourage hint use but may not lead to higher achievement (Aleven et al., 2016). Although hints can reduce floundering, some learners still avoid them (Aleven et al., 2016; Borchers et al., 2025; Li et al., 2024). Success in online learning requires self-regulation and effective help-seeking strategies, yet interventions alone may not overcome avoidance behaviors (Yang, 2023). These differences, however, should be interpreted with caution. The two groups consisted of different students from different schools, and the absence of randomization means that observed differences in behavioral patterns may reflect pre-existing group characteristics rather than the effect of the intervention itself.

3.5 Patterns Strongly Associated with Solved and Unsolved Outcomes for Each Group

In all groups, not skipping was positively associated with solved problems, with the strongest lifts in the without-intervention group (lift = 2.85) and among low LH students (lift = 1.54). Hint use was also positively associated with solved problems in the low LH group (lift = 1.39). Skipping without hints was consistently linked to unsolved outcomes, particularly in the low LH (lift = 1.50) and high LH (lift = 1.48) groups. Both intervention conditions showed skipping behaviors as defining features of unsolved problems. Empirical evidence confirms that help avoidance is a predictor of poor outcomes, while help seeking and persistence foster academic success. Students avoiding assistance show lower performance (Aleven et al., 2006; Li et al., 2024), whereas those who actively seek help cope better with challenges (Li, Che Hassan, & Saharuddin, 2023). Higher interest reduces effort avoidance and promotes persistence (Song, Kim, & Bong, 2019). Even though difficult learning strategies such as interleaved practice are perceived as harder, they lead to better problem-solving ability (Samani and Pan, 2021). Not skipping and using hints are therefore associated with effective self-regulation, while skipping without hints is a consistent marker of unsolved outcomes across all groups.

4. Conclusion and Future Work

This study applied the Apriori algorithm to interaction logs from a mathematics tutoring system to examine behavioral patterns related to LH across differences in LH level, intervention condition, and problem-solving outcome. The analysis found that not using available hints was the most frequent pattern linked to unsolved problems, while persistence behaviors, such as continuing with a problem, appeared less often. Low-LH students showed stronger links between persistence and solved problems, as well as positive associations between hint use and correct solutions. High-LH students displayed more avoidance-oriented patterns, particularly skipping behaviors tied to unsolved results. Students without system-based intervention exhibited stronger persistence-success associations, while those with intervention tended to show more skipping behaviors. Across all groups, continued engagement with problems was consistently related to solved outcomes, whereas avoidance without hint use was more often connected to unsolved ones.

There are several limitations in this study. The dataset did not include demographic or contextual variables that might influence engagement patterns. The LH classification was derived from a Random Forest model trained on behavioral features and grounded in Yates' (2009) validated teacher-rated scale; nonetheless, the use of a model-derived binary label as a proxy for a psychological construct introduces some degree of construct validity limitation in how the low and high LH groupings should be interpreted. The intervention and non-intervention groups were drawn from different schools without randomization, so differences in behavioral patterns between groups cannot be attributed to the intervention and should be interpreted as associations rather than causal effects. The study focused on Grade 8 learners in the Philippines, so patterns may differ in other educational levels or contexts. The session-level unit of analysis does not account for within-student dependencies, as individual students contributed multiple sessions to the dataset. This means that behavioral tendencies of high-frequency contributors are reflected more prominently in the reported patterns, and the findings should be interpreted as characterizing session-level co-occurrences rather than consistent behavioral tendencies within individual learners. Future studies employing student-level aggregation or multilevel modeling approaches could more precisely examine whether the reported associations generalize across individual students. Because behavioral indicators were aggregated at the session level, within-session event sequences were not modeled, and future studies with timestamped action-level logs are encouraged to apply other pattern mining alongside association rule mining. A sensitivity check confirmed that the primary avoidance-related patterns were stable across varied threshold combinations, though future work may explore a wider range of values to further establish robustness. The analysis also relied solely on log data, which captures observable behaviors but not the underlying cognitive or emotional processes that drive them.

Despite these constraints, the observed patterns suggest several opportunities for improving ITS. Adaptive features could be designed to detect early signs of avoidance, such as repeated skipping or low hint use, and to provide timely prompts that encourage persistence and improve help-seeking. For students with higher LH, strategies may involve guiding them toward productive hint use and reinforcing persistence after mistakes. For students with lower LH, approaches could aim to sustain engagement and gradually increase problem difficulty to strengthen resilience. Educators might also use these patterns to identify students who could benefit from targeted support, combining system data with classroom observations for a more complete understanding of learning behaviors. Future enhancements to the AES platform may include real-time pattern detection, dynamic adjustment of problem difficulty, and context-sensitive feedback. Additional studies that integrate behavioral logs with self-reported or interview data could help explain the motivational and emotional factors driving these behaviors.

Ethical Statement: This study used student data collected by the author under ethics clearance from the UE Ethics Review Committee, with informed consent obtained from all participants and full compliance with the Data Privacy Act of 2012.

AI Ethics Statement: The author confirm that they did not use AI when writing this study.

References

Agrawal, Rakesh and Srikant, Ramakrishnan. (1994). Fast Algorithms for Mining Association Rules in Large Databases. *Proceedings of the 20th International Conference on Very Large Data Bases*, pp. 487–499.

Albay, Eduard M. (2020). Towards a 21st Century Mathematics Classroom: Investigating the Effects ofthe Problem-Solving Approach Among Tertiary Education Students. *Asia-Pacific Social Science Review*, 20(2), pp. 8. doi:https://doi.org/10.59588/2350-8329.1303

Aleven, Vincent ; McLaren, Bruce ; Roll, Ido and Koedinger, Kenneth. (2006). Toward Meta-cognitive Tutoring: A Model of Help Seeking with a Cognitive Tutor. *International Journal of Artificial Intelligence in Education*, 16(2), pp. 101–128. doi:10.3233/IRG-2006-16(2)02

Aleven, Vincent ; Roll, Ido ; McLaren, Bruce M and Koedinger, Kenneth R. (2016). Help Helps, But Only So Much: Research on Help Seeking with Intelligent Tutoring Systems. *International Journal of Artificial Intelligence in Education*, 26(1), pp. 205–223. doi:10.1007/s40593-015-0089-1

Amadi, Charity Chizoma ; Agi, W.Comfort and Nwoke, Patience Linus. (2020). Psychological Factors as Correlates of Learned Helplessness in Mathematics among Senior Secondary School Students in Rivers State Nigeria. *International Journal of Innovative Psychology & Social Development*, 8(4), pp. 54–64.

Beck, Joseph E and Gong, Yue. (2013). Wheel-Spinning: Students Who Fail to Master a Skill. In H. Chad Lane ; Kalina Yacef ; Jack Mostow and Philip Pavlik (Eds). *Artificial Intelligence in Education*. Springer Berlin Heidelberg.

Biber, Mahir and Biber, Sezer Köse. (2014). Investigation of the Level of Prospective Teachers' Learned Helplessness in Mathematics in Relation of Various Variables. *Procedia - Social and Behavioral Sciences*, 116, pp. 3484–3488. doi:https://doi.org/10.1016/j.sbspro.2014.01.788

Borchers, Conrad ; Fleischer, Hendrik ; Schanze, Sascha ; Scheiter, Katharina and Aleven, Vincent. (2025). High scaffolding of an unfamiliar strategy improves conceptual learning but reduces enjoyment compared to low scaffolding and strategy freedom. *Computers & Education*, 236, pp. 105364. doi:https://doi.org/10.1016/j.compedu.2025.105364

Bringula, R P ; Alvarez, J N S ; Evangelista, M A C ; So, R B and Gatus, M M. (2015). Technical Description of Equation Sensei: A Mobile-Assisted Learning Application in Mathematics. *2015 IEEE 15th International Conference on Advanced Learning Technologies*, pp. 105–107. doi:10.1109/ICALT.2015.92

Bringula, R P ; Rodrigo, M M T ; Rivera, V ; Malaborbor, R A ; Catacutan-Bangit, A E and Sumi, K. (2025). Pattern of Student Game Setting Selections in an Intelligent Mobile MCSCL using the Apriori Algorithm. *2025 13th International Conference on Information and Education Technology (ICIET)*, pp. 98–102. doi:10.1109/ICIET66371.2025.11046270

Dang, Steven C. and Koedinger, Kenneth R. (2020). The Ebb and Flow of Student Engagement: Measuring motivation through temporal pattern of self-regulation. *Proceedings of The 13th International Conference on Educational Data Mining (EDM 2020)*, pp. 61–68.

De, Sourav ; Dey, Sandip ; Bhatia, Surbhi and Bhattacharyya, Siddhartha. (2022). Chapter 1 - An introduction to data mining in social networks. In Sourav De ; Sandip Dey ; Siddhartha Bhattacharyya and Surbhi B.T. Advanced Data Mining Tools and Methods for Social Computing Bhatia (Eds). *Hybrid Computational Intelligence for Pattern Analysis*. Academic Press. doi:https://doi.org/10.1016/B978-0-32-385708-6.00008-4

Evans, Paul ; Vansteenkiste, Maarten ; Parker, Philip ; Kingsford-Smith, Andrew and Zhou, Sijing. (2024). Cognitive Load Theory and Its Relationships with Motivation: a Self-Determination Theory Perspective. *Educational Psychology Review*, 36(1), pp. 7. doi:10.1007/s10648-023-09841-2

Fancsali, Stephen E. (2014). Causal Discovery with Models: Behavior, Affect, and Learning in Cognitive Tutor Algebra. *Proceedings of the 7th International Conference on Educational Data Mining*, pp. 28–35.

Fu, You ; Ren, Fang and Lin, Jiantao. (2025). Apriori algorithm based prediction of students' mental health risks in the context of artificial intelligence. *Frontiers in Public Health*, Volume 13. doi:https://doi.org/10.3389/fpubh.2025.1533934

Gürefe, Nejla and Bakalım, Orkide. (2018). Mathematics Anxiety, Perceived Mathematics Self-efficacy and Learned Helplessness in Mathematics in Faculty of Education Students. *International Online Journal of Educational Sciences*, 10(3), pp. 147–161. doi:http://dx.doi.org/10.15345/iojes.2018.03.010

Hikmawati, Erna ; Maulidevi, Nur Ulfa and Surendro, Kridanto. (2021). Minimum threshold determination method based on dataset characteristics in association rule mining. *Journal of Big Data*, 8(1), pp. 146. doi:10.1186/s40537-021-00538-3

Hwang, Jihyun. (2019). Relationships among locus of control, learned helpless, and mathematical literacy in PISA 2012: focus on Korea and Finland. *Large-Scale Assessments in Education*, 7(1), pp. 4. doi:10.1186/s40536-019-0072-7

Kuldas, Seffetullah ; Satyen, Lata ; Ismail, Hairul Nizam and Hashim, Shahabuddin. (2014). Greater Cognitive Effort for Better Learning: Tailoring an Instructional Design for Learners with Different Levels of Knowledge and Motivation. *Psychologica Belgica*, 54(4), pp. 350–373. doi:10.5334/pb.aw

Li, Ji ; Fang, Li ; Liu, Yu ; Xie, Jiayu and Wang, Xiaoyu. (2024). Ineffective Learning Behaviors and Their Psychological Mechanisms among Adolescents in Online Learning: A Narrative Review. *Behavioral Sciences (Basel, Switzerland)*, 14(6). doi:10.3390/bs14060477

Li, Ruihua ; Che Hassan, Norlizah and Saharuddin, Norzihani. (2023). College Student's Academic Help-Seeking Behavior: A Systematic Literature Review. *Behavioral Sciences (Basel, Switzerland)*, 13(8). doi:10.3390/bs13080637

Maier, Steven F and Seligman, Martin E P. (2016). Learned helplessness at fifty: Insights from neuroscience. *Psychological Review*, 123(4), pp. 349–367. doi:10.1037/rev0000033

Miranda, J P P and Bringula, R P. (2023). Towards the Development of a Model to Detect Learned Helplessness Among Grade 6 Mathematics Students. *2023 IEEE International Conference on Advanced Learning Technologies (ICALT)*, pp. 373–375. doi:10.1109/ICALT58122.2023.00115

Miranda, John Paul and Bringula, Rex. (2025). Modeling Student Learned Helplessness in Mathematics Using Random Forest. *2025 13th International Conference on Information and Education Technology*, pp. 138–142.

Miranda, John Paul P ; Bringula, Rex P ; Simpao, Laharni S ; Salenga, Jordan L ; Grume, Juvy C ; Nacianceno, Madilaine Claire B ; Loyola, Lester G and Yambao, Jaymark A. (2025). Towards the Development of Detection of Learned Helplessness in Mathematics: Design and Data Collection Challenges from a Developing Country Perspective. In Robert A. Sottilare and Jessica Schwarz (Eds). *Adaptive Instructional Systems*. Springer Nature Switzerland. doi:https://doi.org/10.1007/978-3-031-92970-0_7

Muangprathub, Jirapond ; Boonjing, Veera and Chamnongthai, Kosin. (2020). Learning recommendation with formal concept analysis for intelligent tutoring system. *Heliyon*, 6(10). doi:10.1016/j.heliyon.2020.e05227

Namukasa, Maureen ; Carroll, Meredith ; Marraffino, Matthew D ; Johnson, Cheryl I ; Whitmer, Daphne E ; Igwe, Whitney-Marie and Nash, Morgan. (2023). Not All Pain Leads to Gain: The Role of Learner Engagement in Adaptive Flashcard Training. In Robert A. Sottilare and Jessica Schwarz (Eds). *Adaptive Instructional Systems*. Springer Nature Switzerland. doi:https://doi.org/10.1007/978-3-031-34735-1_2

Nuvvula, Sivakumar. (2016). Learned helplessness. *Contemporary Clinical Dentistry*, 7(4), pp. 426–427. doi:https://doi.org/10.4103/0976-237X.194124

Papadogiannis, Ilias ; Wallace, Manolis and Karountzou, Georgia. (2024). Educational Data Mining: A Foundational Overview. In *Encyclopedia* Vol. 4, Issue 4. doi:10.3390/encyclopedia4040108

Plooy, Eileen du ; Casteleijn, Daleen and Franzsen, Denise. (2024). Personalized adaptive learning in higher education: A scoping review of key characteristics and impact on academic performance and engagement. *Heliyon*, 10(21), pp. e39630. doi:https://doi.org/10.1016/j.heliyon.2024.e39630

Real, E M ; Pimentel, E P and Braga, J C. (2021). Analysis of Learning Behavior in a Programming Course using Process Mining and Sequential Pattern Mining. *2021 IEEE Frontiers in Education Conference (FIE)*, pp. 1–9. doi:10.1109/FIE49875.2021.9637146

Ryan, R M and Deci, E L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), pp. 68–78. doi:https://doi.org/10.1037/0003-066X.55.1.68

Samani, Joshua and Pan, Steven C. (2021). Interleaved practice enhances memory and problem-solving ability in undergraduate physics. *NPJ Science of Learning*, 6(1), pp. 32. doi:10.1038/s41539-021-00110-x

Sideridis, Georgios D. (2003). On the origins of helpless behavior of students with learning disabilities: avoidance motivation? *International Journal of Educational Research*, 39(4), pp. 497–517. doi:https://doi.org/10.1016/j.ijer.2004.06.011

Song, Juyeon ; Kim, Sung-il and Bong, Mimi. (2019). The More Interest, the Less Effort Cost Perception and Effort Avoidance. *Frontiers in Psychology*, Volume 10.

Sowan, Bilal ; Zhang, Li ; Matar, Nasim ; Zraqou, J ; Omar, Firas and Alnatsheh, Athari. (2025). A novel lift adjustment methodology for improving association rule interpretation. *Decision Analytics Journal*, 15, pp. 100582. doi:https://doi.org/10.1016/j.dajour.2025.100582

Spitzer, Markus Wolfgang Hermann and Moeller, Korbinian. (2023). Performance increases in mathematics during COVID-19 pandemic distance learning in Austria: Evidence from an intelligent tutoring system for mathematics. *Trends in Neuroscience and Education*, 31, pp. 100203. doi:https://doi.org/10.1016/j.tine.2023.100203

Su, Man ; Dang, Belle ; Nguyen, Andy and Nagashima, Tomohiro. (2025). Choice-making in an adaptive learning system with motivational pedagogical agents. *NPJ Science of Learning*, 10(1), pp. 77. doi:10.1038/s41539-025-00366-7

Sweller, John. (1988). Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science*, 12(2), pp. 257–285. doi:https://doi.org/10.1207/s15516709cog1202_4

Tang, Zheng ; Jiang, Zhengwei ; Li, Ying ; Yuan, Haowei ; Han, Jiayu and Chen, Chao. (2024). Research on the Association Analysis of Online Learning Behaviors Based on the Apriori Algorithm. *Frontiers in Computing and Intelligent Systems*, 9(2), pp. 18–22. doi:10.54097/4a3h3p03

Tsaparlis, Georgios. (2021). Chapter 5. It Depends on the Problem and on the Solver: An Overview of the Working Memory Overload Hypothesis, Its Applicability and Its Limitations. In *The Royal Society of Chemistry*. doi:https://doi.org/10.1039/9781839163586-00093

Wang, Tengfei ; Xiao, Baorong and Ma, Weixiao. (2022). Student Behavior Data Analysis Based on Association Rule Mining. *International Journal of Computational Intelligence Systems*, 15(1), pp. 32. doi:10.1007/s44196-022-00087-4

Weiner, Bernard. (1986). *An Attributional Theory of Motivation and Emotion* (1st edn). Springer New York, NY. doi:https://doi.org/10.1007/978-1-4612-4948-1

Yang, Albert C M ; Chen, Irene Y L ; Flanagan, Brendan and Ogata, Hiroaki. (2022). How students' self-assessment behavior affects their online learning performance. *Computers and Education: Artificial Intelligence*, 3, pp. 100058. doi:https://doi.org/10.1016/j.caeai.2022.100058

Yang, Fan. (2023). A Systematic Review of Studies Exploring Help-Seeking Strategies in Online Learning Environments. *Online Learning Journal*, 27(1), pp. 107–126. doi:10.24059/olj.v27i1.3400

Yates, Shirley. (2009). Teacher Identification of Student Learned Helplessness in Mathematics. *Mathematics Education Research Journal*, 21(3), pp. 86–106.

Zhang, Yingbin and Paquette, Luc. (2023). Sequential Pattern Mining in Educational Data: The Application Context, Potential, Strengths, and Limitations. In Alejandro Peña-Ayala (Ed.). *Educational Data Science: Essentials, Approaches, and Tendencies: Proactive Education based on Empirical Big Data Evidence*. Singapore : Springer Nature Singapore. doi:10.1007/978-981-99-0026-8_6

Zimmerman, Barry J. (2000). Self-Efficacy: An Essential Motive to Learn. *Contemporary Educational Psychology*, 25(1), pp. 82–91. doi: https://doi.org/10.1006/ceps.1999.1016