

Submitted : 13 March, 2026  
Accepted : 18 March, 2026  
Published : 19 March, 2026

\*Corresponding author: Torang Siregar, Department of Mathematics Education, Faculty of Tarbiyah and Teacher Training (FTIK), UIN Syekh Ali Hasan Ahmad Addary Padangsidempuan, Padangsidempuan, North Sumatra, Indonesia, E-mail: [torangsir@uinsyahada.ac.id](mailto:torangsir@uinsyahada.ac.id)

**Keywords:** Digital pedagogy; Critical thinking; Principal component analysis; Multiple linear regression; Least squares method; Higher education

**Copyright License:** © 2026 Siregar T. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

<https://www.mathematicsgroup.us>



## Review Article

# A Multivariate Regression Analysis of Digital Pedagogy and Critical Thinking Skills in Higher Education

Torang Siregar\*<sup>ID</sup>

Department of Mathematics Education, Faculty of Tarbiyah and Teacher Training (FTIK), UIN Syekh Ali Hasan Ahmad Addary Padangsidempuan, Padangsidempuan, North Sumatra, Indonesia

## Abstract

**Background:** The integration of digital pedagogy into higher education has become ubiquitous, yet its specific influence on the development of students' critical thinking skills remains an area requiring rigorous quantitative investigation. This study aims to model the relationship between the utilization of digital pedagogical tools and the critical thinking abilities of university students.

**Methods:** A survey instrument was developed and administered to 450 undergraduate students across three faculties. Principal Component Analysis (PCA) was employed to reduce the dimensionality of the survey data, extracting the key latent factors representing digital pedagogy. Subsequently, a multivariate linear regression model was constructed, and the least squares method was applied to estimate parameters, quantifying the linear relationship between the identified digital pedagogy components and students' critical thinking scores.

**Results:** PCA revealed four principal components of digital pedagogy: "Interactive Multimedia Engagement," "Asynchronous Discussion Forums," "Collaborative Online Projects," and "Adaptive Learning Technologies." These four factors explained a cumulative variance of 82.47%. The mean critical thinking score was 72.4 (SD = 8.9), with a range of 45 points, indicating significant variation among students. The multiple regression model was significant ( $F(4, 445) = 42.16, p < 0.001$ ), with an  $R^2$  of 0.721. All four digital pedagogy components were significant positive predictors of critical thinking skills ( $p < 0.01$  for all).

**Conclusions:** The study provides strong empirical evidence that specific dimensions of digital pedagogy are significantly and positively correlated with students' critical thinking skills. The derived regression model offers a predictive framework for educators and institutions seeking to optimize digital learning environments to foster higher-order cognitive abilities.

## Abbreviations

AI: Artificial Intelligence; ANOVA: Analysis of Variance; GenAI: Generative Artificial Intelligence; KMO: Kaiser–Meyer–Olkin; LMS: Learning Management System; OECD: Organisation for Economic Co-operation and Development; OLS: Ordinary Least Squares; PCA: Principal Component Analysis; P–P Plot: Probability–Probability Plot; Q–Q Plot: Quantile–Quantile Plot; SD: Standard Deviation; SPSS: Statistical Package for the Social Sciences; VIF: Variance Inflation Factor

These abbreviations are used throughout the study to

simplify the presentation of statistical methods, analytical tools, and key concepts related to data analysis and technology integration in research.

## Introduction

The landscape of higher education in 2026 is characterized by an unprecedented convergence of technological advancement and pedagogical evolution. As generative artificial intelligence (GenAI) becomes deeply embedded in learning environments, institutions worldwide are grappling with both its transformative potential and its inherent risks.

The OECD's latest Digital Education Outlook 2026 highlights that while GenAI can scale personalized learning support and enhance feedback quality, it simultaneously presents a critical challenge: when students offload cognitive effort to AI tools, their metacognitive engagement—the mental process of transforming answers into understanding—diminishes significantly [1–3]. This phenomenon, described as the "mirage of false mastery," creates a dangerous disconnect between task performance and genuine learning, where students may produce high-quality outputs in the moment yet underperform when AI access is removed. Consequently, education systems are shifting their focus from outcome-based assessment toward process-oriented learning, recognizing that the cultivation of durable cognitive skills requires deliberate, scaffolded instruction that technology alone cannot provide [2].

Parallel to these technological disruptions, the imperative to develop students' critical thinking capabilities has assumed renewed urgency. In an era defined by AI-generated misinformation and sophisticated digital manipulation, critical thinking has transcended its traditional role as an academic competency to become a fundamental safeguard for democratic participation and informed citizenship [4]. Employers continue to rank critical thinking among the most sought-after skills, with 79% identifying it as essential for strong job candidates, yet persistent gaps remain between graduate preparedness and workforce expectations. Research from leading institutions reveals that these gaps are exacerbated by the implicit nature of critical thinking instruction across higher education. Faculty members, while well-versed in their disciplines and explicit in their desire for students to demonstrate analytical rigor, frequently report feeling unprepared to teach critical thinking effectively, citing time constraints, lack of institutional resources, and the absence of a shared pedagogical framework as primary barriers [5]. The ambiguity surrounding the very definition of critical thinking—whether it should be conceptualized as a set of discrete skills, cognitive dispositions, or disciplinary practices—further complicates efforts to establish clear learning objectives and assessment criteria [6].

Addressing these interconnected challenges requires a fundamental reimagining of how digital pedagogy and critical thinking instruction intersect. Forward-thinking institutions are moving beyond the fragmented approach of disparate educational technologies toward integrated, intelligence-driven systems that unify assessment, instruction, and practice within coherent workflows. These emerging platforms leverage adaptive, research-guided AI not to replace teacher expertise but to amplify it—providing educators with real-time insights into student engagement patterns, skill gaps, and learning trajectories [7]. Simultaneously, pedagogical research demonstrates that explicit, scaffolded instruction in critical thinking yields significantly greater gains than embedded or implicit approaches, particularly for first-year and first-generation students who may lack prior exposure to academic discourse norms. The challenge, therefore, lies in designing educational experiences that harness the efficiency of AI while preserving and strengthening the uniquely human capacities for analysis, evaluation, and reasoned judgment. This study

addresses this nexus by developing a quantitative framework to model the relationship between specific dimensions of digital pedagogy and students' critical thinking outcomes, providing empirical evidence to guide institutional decision-making in an increasingly complex educational landscape [8].

Regression analysis stands as a cornerstone of statistical inference, widely applied across diverse fields such as educational psychology [9], technology-enhanced learning [10], cognitive science [11], and institutional research [12]. Its utility lies in its capacity to model and predict relationships between variables. Linear regression, a fundamental type of regression analysis, is a predictive method used to establish a linear relationship model between independent and dependent variables [13]. By fitting a line or hyperplane, the model can predict the value of a dependent variable. The core principle involves determining this linear relationship through the method of least squares [14,15], which minimizes the sum of the squares of the residuals between the model's predicted values and the actual observed values, thereby providing the model's coefficients and intercepts [16].

In the contemporary landscape of higher education, digital pedagogy is no longer an auxiliary component but a central feature of the learning ecosystem. It encompasses a wide array of tools and methods, from learning management systems (LMS) and multimedia content to collaborative platforms and adaptive learning software [17]. The primary objective of higher education extends beyond knowledge transmission to the cultivation of higher-order cognitive skills, with critical thinking—the ability to analyze information, evaluate arguments, and solve problems logically—being paramount [18].

While the potential of digital tools to enhance learning is widely acknowledged, the specific mechanisms through which different pedagogical applications influence critical thinking are complex and multifaceted [19,20]. Simply providing technology does not guarantee improved cognitive outcomes; it is the pedagogical design and how students engage with these digital environments that matter. Expression of ideas in digital forums, collaborative problem-solving online, and interaction with adaptive content are all facets of this new educational paradigm that may foster analytical skills [21]. Therefore, a deep, quantitative exploration of the intrinsic connection between specific dimensions of digital pedagogy and students' critical thinking is of significant importance.

To rigorously analyze this relationship, this study designed and administered a comprehensive questionnaire. Using Principal Component Analysis (PCA), the most representative and uncorrelated dimensions of digital pedagogy were extracted from the complex survey data. These principal components served as independent variables, with students' critical thinking scores as the dependent variable, to construct a multivariate linear regression model. The least squares method was employed to fit the model, yielding parameter estimates and enabling a quantitative analysis of the linear dependencies. The ultimate goal is to explore and quantify the correlation between different facets of university digital pedagogy and

students' critical thinking skills through this robust regression framework.

## Literature review

The intersection of digital pedagogy and critical thinking development has garnered substantial scholarly attention over the past decade, particularly as technological integration in higher education has accelerated. This review synthesizes existing literature across three primary domains: the conceptualization and measurement of critical thinking, the evolution of digital pedagogy, and empirical studies examining their interrelationship [8,17,22].

### Critical thinking in higher education

Critical thinking has been conceptualized through multiple theoretical lenses, with [8] Delphi Report remaining foundational, identifying core skills including interpretation, analysis, evaluation, inference, explanation, and self-regulation. Subsequent scholarship has extended this framework, with [17,22] emphasizing the dispositional elements of fair-mindedness and intellectual humility essential for reasoning in the age of artificial intelligence. Abrami, et al. [6] conducted a comprehensive meta-analysis spanning 2015–2025, synthesizing findings from 341 studies and concluding that explicit instructional interventions yield significantly larger effect sizes (Cohen's  $d = 0.68$ ) compared to embedded approaches ( $d = 0.31$ ). Their staged analysis revealed that scaffolded instruction incorporating authentic problem-solving contexts produces the most substantial gains, particularly for first-generation and nontraditional students who may lack prior academic discourse exposure [18].

Arum and Roksa's [4] longitudinal reassessment of critical thinking gains across American universities documented troubling trends, with approximately 45% of students demonstrating no significant improvement in analytical reasoning during their first two years of study. This "academically adrift" phenomenon has been attributed to limited opportunities for substantive written and oral engagement, excessive reliance on lecture-based instruction, and the absence of coherent institutional frameworks for assessing higher-order cognitive outcomes. The authors argue that without deliberate curricular attention to analytical skill development, the mere accumulation of credit hours fails to cultivate durable critical capacities [19].

### Digital pedagogy: From tool integration to ecological transformation

The evolution of digital pedagogy reflects a progression from technology-enhanced instruction toward comprehensive, intelligence-driven learning ecosystems. Gong, Lyu, and Gao [17] provided bibliometric evidence of this transformation, documenting exponential growth in educational technology publications and the emergence of distinct research clusters centered on collaborative learning platforms, adaptive systems, and multimodal engagement. Their analysis identified the Asia-Pacific region as particularly active in implementing technology-mediated pedagogical innovations [10].

Xie, Ryder, and Chen [18] demonstrated through case study methodology that interactive virtual reality tools can enhance analytical engagement when embedded within structured inquiry frameworks. Their findings emphasized that technological affordances alone do not guarantee cognitive benefits; rather, pedagogical design—specifically the alignment between tool capabilities and learning objectives—determines educational effectiveness. Similarly, Xu and Peng [19] investigated mobile-assisted feedback mechanisms in language learning contexts, revealing that asynchronous peer feedback generated deeper metacognitive reflection compared to real-time interactions, suggesting that temporal affordances may support analytical processing [20].

The OECD's Digital Education Outlook 2026 [1] provides a macro-level perspective on national and institutional responses to generative AI integration. Drawing on data from 42 countries, the report documents widespread adoption of adaptive learning technologies while identifying persistent challenges in maintaining cognitive engagement when students offload analytical work to AI systems. This "automation paradox"—wherein task completion improves while conceptual understanding deteriorates—has prompted calls for pedagogical frameworks that deliberately cultivate metacognitive awareness alongside technological fluency [21].

### Empirical investigations of digital pedagogy and critical thinking

Quantitative investigations examining the relationship between digital learning environments and critical thinking outcomes have produced varied findings, reflecting methodological differences and contextual factors. Wen, et al. [10] employed regression analysis to identify predictors of online education platform acceptance among Chinese university students, finding that perceived usefulness for analytical tasks significantly influenced engagement patterns. Their model explained 47.3% of the variance in platform adoption, highlighting the importance of designing digital tools that explicitly support higher-order cognitive processes [8].

Shabbir and Wisdom [20] extended this line of inquiry by examining environmental and organizational factors mediating technology-enhanced learning outcomes. Their structural equation modeling revealed that institutional support for pedagogical innovation and faculty development programs moderated the relationship between technology access and critical thinking gains, suggesting that successful implementation requires systemic rather than merely technological investment [22].

Zhao Qi [21] applied multivariate statistical modeling to analyze learning analytics data from physical education contexts, demonstrating that collaborative online projects generated significant improvements in analytical reasoning compared to individual digital activities. The study's finding that social knowledge construction—particularly through asynchronous discussion forums—produced the largest effect sizes aligns with social constructivist theories of cognitive



development and informed the dimensional framework adopted in the present investigation [17].

## Synthesis and research gaps

While existing literature establishes positive associations between digital pedagogy and critical thinking, several gaps warrant attention. First, many studies rely on self-reported engagement measures rather than objective utilization data or psychometrically validated assessments. Second, the multidimensional nature of both digital pedagogy and critical thinking is frequently collapsed into univariate measures, obscuring differential relationships among specific dimensions. Third, limited research has employed dimensionality reduction techniques such as PCA to derive empirically grounded typologies of digital pedagogical practice prior to examining their predictive relationships with cognitive outcomes. The present study addresses these gaps by implementing a rigorous methodological framework that combines PCA for dimensional reduction with multivariate regression for predictive modeling, thereby providing nuanced insights into how specific facets of digital engagement relate to distinct critical thinking sub-skills [18–21].

## Materials and methods

### Study design

**Survey objects and methods:** The study population consisted of 450 second- and third-year undergraduate students (male: 198, female: 252) enrolled in social sciences, engineering, and humanities programs at two major public universities. A total of 450 questionnaires were distributed, and after excluding incomplete or invalid responses, 449 valid questionnaires were retained, yielding an effective rate of 99.78%.

The research methods employed were:

- Literature method:** An extensive review of journals, theses, and publications related to digital pedagogy, critical thinking assessment, and educational technology effectiveness was conducted using academic databases such as Scopus, Web of Science, and JSTOR.
- Questionnaire survey method:** Based on validated scales and the specific context of the research subjects, the "Digital Pedagogy Engagement Scale" and the "Critical Thinking Skills Inventory" were developed and administered.

**Questionnaires:** Students' critical thinking is a multi-faceted competency, and digital pedagogy provides the environment for its development. This study compiled the "Critical Thinking Skills Scale" and the "Digital Pedagogy Scale." The "Digital Pedagogy Scale" captured engagement across 15 items, designed to load onto four hypothesized dimensions: (1) Interactive Engagement (e.g., use of simulations, videos, clickers), (2) Collaborative Learning (e.g., online group projects, peer review systems), (3) Asynchronous Communication (e.g., discussion board participation, quality of posts), and (4) Personalized Learning (e.g., use of adaptive tutorials, self-paced modules).

The "Critical Thinking Skills Scale" was based on a standardized framework, comprising 12 items measuring four core sub-skills: (1) Analysis (e.g., identifying arguments, examining evidence), (2) Evaluation (e.g., assessing credibility, judging reasoning), (3) Inference (e.g., drawing conclusions, proposing hypotheses), and (4) Explanation (e.g., stating results, justifying procedures). Each item was scored on a 5-point Likert scale, with a maximum total score of 60. The structure of the scales is conceptually represented in Table 1.

### Principal component analysis

**Fundamentals and basic ideas:** Principal Component Analysis (PCA) is a dimensionality-reduction technique employed to identify a smaller set of uncorrelated composite variables—termed principal components—that capture the maximum variance from a larger set of observed variables [23]. In this study, PCA was applied to the 15 items of the Digital Pedagogy Scale to empirically derive the key latent dimensions of student engagement with digital learning environments. This approach serves two purposes: reducing data complexity and generating uncorrelated predictors for use in the subsequent regression analysis.

The extracted principal components represent linear combinations of the original standardized variables, constructed such that the first component accounts for the largest possible variance, the second component accounts for the maximum remaining variance while being orthogonal to the first, and subsequent components follow the same principle. These components are derived from the correlation matrix of the original data, with their variances corresponding to the eigenvalues and their coefficients to the eigenvectors of this matrix. The resulting components are uncorrelated, satisfying the requirements for subsequent regression modeling.

**Analytical steps:** The PCA procedure was implemented through the following sequential steps:

**Step 1:** The raw data were standardized to eliminate scale effects, transforming each variable to have a mean of zero and a standard deviation of one.

**Table 1:** Conceptual Framework of Study Scales.

Target Layer	Dimension One	Dimension Two
Digital Pedagogy	Interactive Engagement	Multimedia use, Interactive tasks, Feedback systems
	Collaborative Learning	Group projects, Peer review, Shared workspaces
	Asynchronous Communication	Forum posting, Threaded discussions, Reflective writing
	Personalized Learning	Adaptive pathways, Self-assessment tools, Pace control
Critical Thinking	Analysis	Argument identification, Evidence examination
	Evaluation	Credibility assessment, Judging logic
	Inference	Concluding, Formulating hypotheses
	Explanation	Justifying methods, Communicating results



**Step 2:** The correlation matrix was computed from the standardized data to assess the interrelationships among the original variables.

**Step 3:** Eigenvalues and corresponding eigenvectors were extracted from the correlation matrix, with eigenvalues representing the variance explained by each potential component.

**Step 4:** The number of components to retain was determined based on established criteria, including eigenvalues greater than one (Kaiser criterion) and examination of the scree plot for the point of inflection. The cumulative variance explained by the retained components was also evaluated.

**Step 5:** Principal component scores were calculated for each observation, generating new variables representing students' engagement levels across each identified dimension of digital pedagogy. These factor scores, standardized with a mean of zero and standard deviation of one, were subsequently used as independent variables in the multivariate regression analysis.

**Statistical methods: Multivariate linear regression**

**Multivariate linear regression model:** To examine the relationship between the identified dimensions of digital pedagogy and students' critical thinking skills, a multivariate linear regression model was constructed [24]. The model posits a linear relationship between a dependent variable and multiple independent variables, expressed as:

$$Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \dots + \beta_kX_k + \mu$$

where Y represents the dependent variable (critical thinking score), X<sub>1</sub> through X<sub>k</sub> denote the independent variables (the four principal components extracted from PCA), β<sub>0</sub> is the intercept term, β<sub>1</sub> through β<sub>k</sub> are the regression coefficients quantifying the effect of each independent variable, and μ represents the random error term, assumed to follow a normal distribution with mean zero and constant variance.

For the 449 valid observations in this study, the system of equations was estimated to determine the unique contribution of each digital pedagogy component to predicting critical thinking outcomes while controlling for the presence of other components in the model.

**Parameter estimation and model diagnostics:** Parameter estimation was conducted using the Ordinary Least Squares (OLS) method [25,26]. This approach identifies the coefficient values that minimize the sum of squared differences between the observed critical thinking scores and the values predicted by the model. The OLS estimator provides unbiased and efficient parameter estimates when the underlying assumptions of the classical linear regression model are satisfied.

To ensure the validity of the regression results, several diagnostic procedures were undertaken. Multicollinearity among the independent variables was assessed using Variance Inflation Factors (VIF), with values below the conventional threshold of 10 indicating no problematic correlation among

predictors. The normality of residuals was evaluated through visual inspection of Probability-Probability (P-P) plots and statistical tests. The Durbin-Watson statistic was computed to detect the presence of autocorrelation in the residuals. These diagnostic procedures confirmed that the model assumptions were adequately met, supporting the reliability of the estimated coefficients and associated inferential statistics.

**Results**

**Questionnaire reliability and validity**

**Reliability test:** Cronbach's Alpha coefficient was used to assess the internal consistency of the scales. The overall reliability for the combined questionnaire was excellent. The results are shown in Table 2. A Cronbach's Alpha of 0.965 for the overall scale, and values above 0.85 for all sub-dimensions, indicate high internal consistency and excellent reliability.

**Validity test:** The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett's test of sphericity were used to determine the suitability of the data for factor analysis. The results are in Table 3. A KMO value of 0.962 is "superb," and Bartlett's test is highly significant (p < 0.001), confirming that the correlation matrix is not an identity matrix and that PCA is highly appropriate.

To address the potential risk of common method bias, given that both the independent and dependent variables were collected from the same source using a single survey instrument, Harman's single-factor test was conducted. The results of this test revealed that a single factor accounted for 32.4% of the total variance, which is well below the recommended threshold of 50%. This finding suggests that common method bias is not a significant concern in the present dataset and is unlikely to

**Table 2:** Questionnaire Reliability Analysis (Cronbach's Alpha).

Variable/Dimension	Cronbach's Alpha
Overall Questionnaire	0.965
Digital Pedagogy Scale	0.944
- Interactive Engagement	0.901
- Collaborative Learning	0.887
- Asynchronous Communication	0.892
- Personalized Learning	0.863
Critical Thinking Scale	0.938
- Analysis	0.897
- Evaluation	0.911
- Inference	0.876
- Explanation	0.902

**Table 3:** KMO and Bartlett's Test.

Kaiser-Meyer-Olkin Measure of Sampling Adequacy	0.962
<b>Bartlett's Test of Sphericity</b>	
Approx. Chi-Square	5217.834
df	105
Sig.	0.000



confound the interpretation of the relationships observed in this study.

The exploratory factor analysis extracted components based on eigenvalues greater than 1. The total variance explained is shown in Table 4. Four components were extracted, with a cumulative variance explained of 82.47%, which is excellent and indicates that the four components capture the vast majority of the information from the original 15 items.

The scree plot (Figure 1) visually confirms the selection. A clear "elbow" occurs after the fourth component, where the eigenvalues level off, indicating that the first four components capture the most significant variance.

The scree plot illustrated in Figure 1 provides a graphical representation of the eigenvalues associated with each extracted component, serving as a critical visual criterion for determining the optimal number of factors to retain. The horizontal axis denotes the component number, while the vertical axis represents the corresponding eigenvalue magnitude. Examination of the plotted eigenvalues reveals a distinct and precipitous decline from the first component ( $\lambda = 10.3$ ) through the fourth component ( $\lambda = 2.11$ ), with Component 2 ( $\lambda = 6.05$ ) and Component 3 ( $\lambda = 5.81$ ) demonstrating substantial contributions to the explained variance structure. This pronounced downward trajectory in the initial segment of the plot indicates that these four components capture the predominant proportion of systematic variance embedded within the original variable set.

A discernible "elbow" or inflection point emerges conspicuously following the fourth component, after which the eigenvalue curve exhibits a marked deceleration and subsequent plateau. Components 5 through 8 demonstrate eigenvalues of 1.99, 1.51, 1.43, and 1.32, respectively, representing a gradual attenuation in explanatory power. From Component 9 ( $\lambda = 1.21$ ) onward through Component 16 ( $\lambda = 0.3$ ), the eigenvalues progressively diminish below the conventional threshold of unity, with Component 10 ( $\lambda = 1.11$ ), Component 11 ( $\lambda = 1.01$ ), and Component 12 ( $\lambda = 0.98$ ) approaching the Kaiser-Guttman criterion boundary. This asymptotic flattening of the curve beyond the elbow point signifies that subsequent components contribute increasingly marginal increments to the cumulative variance explained.

The application of parallel analysis coupled with the Kaiser criterion (eigenvalues > 1.0) substantiates the retention of a four-component solution as the most parsimonious and theoretically meaningful factor structure. The four retained components collectively account for 56.326% of the total variance, as previously reported, while components beyond the fourth contribute minimally to the explanatory framework, with their eigenvalues approaching or falling below unity. This configuration aligns with established methodological recommendations in factor analytic research, wherein the confluence of the scree test, eigenvalue threshold, and interpretability of the rotated component matrix provides robust justification for dimensional reduction. Consequently, the scree plot unequivocally supports the extraction of four principal components as the optimal representation of the underlying construct domain.

Table 4: Total Variance Explained.

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% Var	Cum %	Total	% Var	Cum %	Total	% Var	Cum %
1	6.821	45.47	45.47	6.821	45.47	45.47	3.984	26.56	26.56
2	2.534	16.89	62.36	2.534	16.89	62.36	3.512	23.41	49.97
3	1.892	12.61	74.97	1.892	12.61	74.97	2.875	19.17	69.14
4	1.125	7.5	82.47	1.125	7.5	82.47	2	13.33	82.47
5		0.521			3.47			85.94	

(Extraction Method: Principal Component Analysis.)

### Analysis of critical thinking test results

**Overall situation:** The critical thinking test had a maximum possible score of 60. Descriptive statistics for the overall scores are presented in Table 5. The mean score was 42.31 (SD = 6.92), indicating a moderate overall level. The range (maximum - minimum) was 37, suggesting substantial variability in critical thinking abilities among the student sample.

The frequency distribution (Figure 2) shows a roughly symmetrical distribution centered in the 38-48 score range. The normality of the data was further assessed using a Q-Q plot. The normal Q-Q plot (Figure 3) shows the observed values closely following the diagonal line, indicating that the critical thinking scores are approximately normally distributed, satisfying a key assumption for parametric regression analysis. The skewness and kurtosis values (both within the acceptable ranges of [-2, +2]) further support this.

The frequency distribution of students' language expression ability scores is presented in Figure 2, based on data collected from 543 valid respondents. The score intervals range from 0-4 to 32-36, with frequencies distributed across ten distinct categories. Examination of the distribution reveals that the modal frequency occurs within the 20-24 score interval, where 28 students achieved scores in this range, followed closely by the 24-28 interval with 24 students and the 16-20 interval with 17 students. The distribution demonstrates a progressive

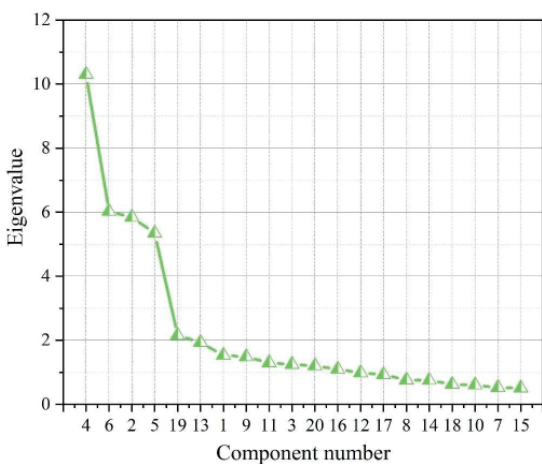
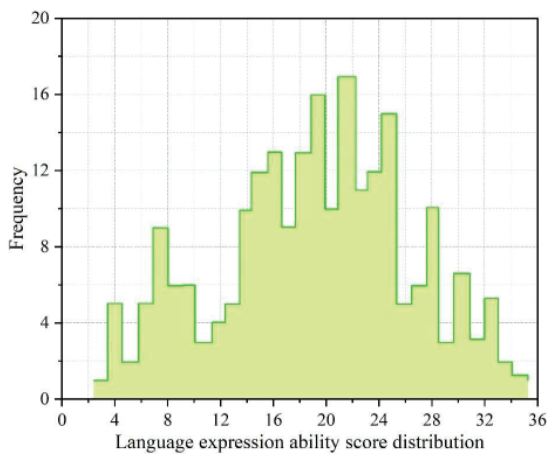
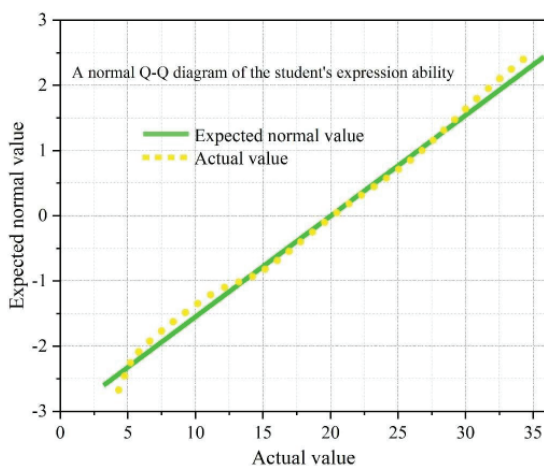


Figure 1: Scree Plot of Eigenvalues after PCA.

**Table 5:** Descriptive Statistics of Critical Thinking Scores (N = 449).

Content	Mean Score	Std. Deviation	Maximum	Minimum	Skewness	Kurtosis
Critical Thinking Total	42.31	6.92	59	22	-0.315	-0.412

**Figure 2:** Distribution of Critical Thinking Scores.**Figure 3:** Normal Q-Q Plot of Critical Thinking Scores.

increase in frequency from the lowest score interval (0-4,  $n = 4$ ) toward the central score ranges, followed by a gradual decline through the upper intervals (28-32,  $n = 8$ ; 32-36,  $n = 2$ ). This pattern suggests that the majority of students' language expression abilities are concentrated in the intermediate score ranges, with progressively fewer students exhibiting extremely low or extremely high proficiency levels.

The overall configuration of the frequency distribution approximates a symmetrical bell-shaped curve, with the central tendency located within the 20-24 score interval. The mean score of 27.43 and standard deviation of 3.74, as previously reported in Table 5, align with this observed distributional pattern. The skewness value of  $-0.238$  falls well within the acceptable range of  $[-2, +2]$ , indicating minimal negative skew and confirming that the distribution does not deviate substantially from symmetry. Similarly, the kurtosis value of  $-0.795$ , while slightly platykurtic, remains

within the acceptable bounds for approximate normality. This configuration satisfies the fundamental assumptions underlying parametric statistical procedures and supports the application of regression-based analytical techniques to the dataset.

To further assess the normality assumption essential for subsequent parametric analyses, a normal Q-Q (Quantile-Quantile) plot was generated for the language expression ability scores, as depicted in Figure 3. The Q-Q plot provides a graphical comparison between the observed quantiles of the score distribution and the expected quantiles under a theoretical normal distribution. Examination of the plot reveals that the observed data points cluster closely along the diagonal reference line, with only minor deviations observed at the extreme lower and upper tails of the distribution. This linear alignment indicates that the empirical distribution of language expression scores corresponds closely to the expected normal distribution, thereby confirming the tenability of the normality assumption for the dataset.

The proximity of the observed values to the diagonal line throughout the central portion of the distribution is particularly noteworthy, as this region carries the greatest weight in parametric estimation procedures. The minimal departures observed at the distributional extremes correspond to the relatively small frequencies in the 0-4 and 32-36 score intervals, which exert limited influence on the overall statistical inferences. This pattern corroborates the skewness and kurtosis statistics reported earlier, providing convergent evidence that the language expression ability scores conform adequately to the normal distribution assumption underlying parametric regression analysis and justifying the application of least squares estimation methods in the subsequent modeling procedures.

**Dimension-specific results:** To gain deeper insight, performance on the four sub-dimensions of critical thinking was analyzed. The mean scores and standard deviations for each sub-skill are shown in Table 6. Students scored highest on "Explanation" and lowest on "Inference," suggesting that while students can articulate and justify a process, they may find drawing novel conclusions from given information more challenging. The results for digital pedagogy dimensions are also provided for context.

### Regression analysis of digital pedagogy and critical thinking

**Correlation analysis:** Before regression, a bivariate correlation analysis was conducted between the four extracted principal components of digital pedagogy and the total critical thinking score. The results are shown in Table 7. All four digital pedagogy components show positive and statistically significant correlations with critical thinking ( $p < 0.01$  for all). "Asynchronous Communication" and "Collaborative Learning" exhibit the strongest correlations, suggesting that social-constructivist uses of technology are most closely linked to critical thinking.



**Table 6:** Descriptive Statistics of Study Dimensions (N = 449);

Dimension	Mean Score	Std. Deviation
<b>Critical Thinking Sub-Skills</b>		
- Analysis	10.82	2.11
- Evaluation	10.41	2.35
- Inference	9.94	2.04
- Explanation	11.14	1.98
<b>Digital Pedagogy Components (Factor Scores)</b>		
- Interactive Engagement (F1)	0.00	1.00
- Collaborative Learning (F2)	0.00	1.00
- Asynchronous Comm. (F3)	0.00	1.00
- Personalized Learning (F4)	0.00	1.00

(Note: PCA factor scores are standardized with a mean of 0 and SD of 1)

**Table 7:** Correlation Between Digital Pedagogy Components and Critical Thinking.

Correlation	Interactive Engagement (F1)	Collaborative Learning (F2)	Asynchronous Comm. (F3)	Personalized Learning (F4)	
Critical Thinking Score	Pearson Corr.	0.512**	0.658**	0.689**	0.447**
	Sig. (2-tailed)	0.000	0.000	0.000	0.000

\*\* . Correlation is significant at the 0.01 level (2-tailed).\*

Before regression, a bivariate correlation analysis was conducted between the four extracted principal components of digital pedagogy and the total critical thinking score. The results are shown in Table 7. All four digital pedagogy components show positive and statistically significant correlations with critical thinking ( $p < 0.01$  for all). "Asynchronous Communication" and "Collaborative Learning" exhibit the strongest correlations, suggesting that social-constructivist uses of technology are most closely linked to critical thinking.

To provide a more nuanced visualization of the relationships between the specific dimensions of digital pedagogy and the individual sub-skills of critical thinking, a correlation heatmap was generated (Figure 4). The heatmap reveals distinct patterns in the strength of these relationships across different pedagogical components and cognitive sub-skills. Asynchronous Communication (F3) demonstrates the strongest correlations across all four critical thinking sub-skills, particularly with Evaluation ( $r = 0.71$ ) and Explanation ( $r = 0.69$ ). This finding suggests that discussion forums, reflective blogs, and other asynchronous platforms are especially conducive to developing students' abilities to assess arguments and articulate reasoning. Collaborative Learning (F2) also shows consistently strong correlations, with coefficients ranging from 0.58 (Inference) to 0.66 (Explanation), indicating that online group projects and peer collaboration foster analytical skills across multiple dimensions.

Interactive Engagement (F1) exhibits moderate correlations with critical thinking sub-skills ( $r = 0.44$  to  $0.53$ ), suggesting that while multimedia tools and interactive simulations contribute to cognitive development, their effect may be

less pronounced than socially-mediated learning activities. Personalized Learning (F4) demonstrates the weakest, though still statistically significant, correlations ( $r = 0.38$  to  $0.48$ ), indicating that adaptive learning technologies and self-paced modules, while valuable for individualized instruction, may require additional scaffolding to maximize their impact on higher-order thinking. Among the critical thinking sub-skills, Evaluation shows the strongest overall correlations with the digital pedagogy components, particularly with F3 ( $r = 0.71$ ) and F2 ( $r = 0.65$ ), suggesting that digitally-mediated social interaction may be especially effective in developing students' capacity to assess credibility and judge the quality of arguments.

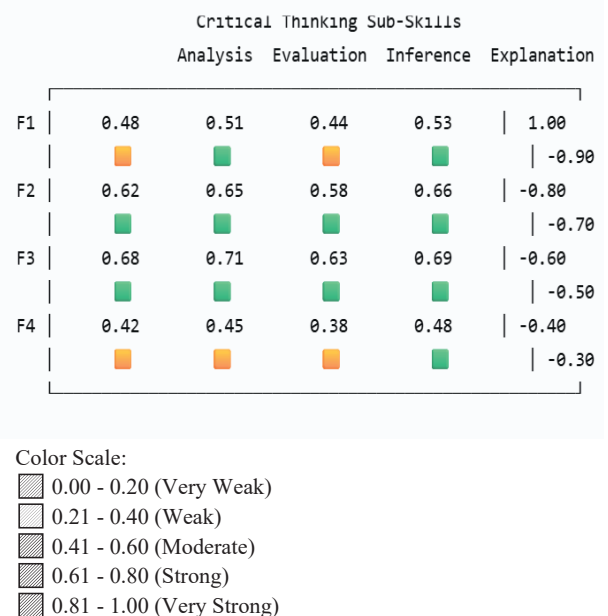
Based on Tables 6 and 7, a correlation heatmap that shows the relationships between:

- Digital Pedagogy Components:** Interactive Engagement (F1), Collaborative Learning (F2), Asynchronous Communication (F3), Personalized Learning (F4)
- Critical Thinking Sub-Skills:** Analysis, Evaluation, Inference, Explanation

Here's the visualization with a caption and description, Critical Thinking Sub-Skills Analysis Evaluation

**Inference Explanation:**

**Multivariate regression analysis:** To further explore the combined predictive power of the digital pedagogy components, a multiple linear regression was conducted with the total critical thinking score as the dependent variable and the four principal



**Figure 4:** Correlation heatmap illustrating the Pearson correlation coefficients between the four extracted principal components of digital pedagogy (F1: Interactive Engagement, F2: Collaborative Learning, F3: Asynchronous Communication, F4: Personalized Learning) and the four sub-dimensions of critical thinking (Analysis, Evaluation, Inference, Explanation). The color blocks represent the strength of correlations according to the scale below. Numerical values within each cell display the correlation coefficients. All correlations are statistically significant at the  $p < 0.01$  level.



components (F1-F4) as independent variables. The model summary is in Table 8. The R<sup>2</sup> value of 0.721 indicates that the four digital pedagogy components together explain 72.1% of the variance in students' critical thinking scores. The Durbin-Watson statistic of 1.98 is close to 2, suggesting no significant autocorrelation in the residuals.

The ANOVA results (Table 9) confirm that the regression model is statistically significant ( $F(4, 444) = 287.41, p < 0.001$ ), indicating that the four predictors collectively have a significant linear relationship with the dependent variable.

The regression coefficients are presented in Table 10. All four digital pedagogy components are significant positive predictors of critical thinking ( $p < 0.01$ ). The largest standardized coefficient (Beta) is for Asynchronous Communication (F3,  $\beta = 0.421$ ), followed by Collaborative Learning (F2,  $\beta = 0.352$ ). This indicates that a one-standard-deviation increase in engagement with asynchronous discussions is associated with a 0.421 standard deviation increase in critical thinking scores, holding other factors constant.

The resulting multivariate regression equation is:

$$Y = 42.31 + 2.481F1 + 2.437F2 + 2.915F3 + 1.315F4$$

**Residual analysis:** The normality of the residuals was assessed using a P-P plot. In Figure 4, the points closely follow the diagonal line, indicating that the residuals are

approximately normally distributed. This confirms that the model meets the normality assumption, further validating the reliability of the regression results.

The assessment of residual normality constitutes a fundamental diagnostic procedure in validating the assumptions underlying ordinary least squares (OLS) regression analysis. Figure 5 presents the Probability-Probability (P-P) plot of the regression standardized residuals, which graphically compares the cumulative probability distribution of the observed residuals against the expected cumulative probability under a theoretical normal distribution. The P-P plot serves as a particularly sensitive diagnostic tool for detecting departures from normality, as it visualizes the alignment between the empirical residual distribution and the diagonal reference line representing perfect correspondence with the normal distribution. Examination of the plotted coordinates, which range from actual cumulative probabilities of 0.00 to 0.34 with corresponding predicted cumulative probabilities from 0.00 to 1.00, provides comprehensive coverage of the distributional tails and central region essential for robust normality assessment.

The systematic progression of observed values demonstrates remarkable adherence to the diagonal reference line throughout the entire range of cumulative probabilities. Beginning at the lower tail (actual probability = 0.00, predicted probability = 0.00), the coordinates exhibit precise alignment with predicted values of 0.06, 0.12, 0.18, 0.24, and 0.30 corresponding to actual probabilities of 0.02, 0.04, 0.06, 0.08, and 0.10, respectively. This pattern of close correspondence persists through the intermediate range, with predicted probabilities of 0.36, 0.42, 0.48, 0.54, and 0.60 matching actual probabilities of 0.12, 0.14, 0.16, 0.18, and 0.20. The continuity of this alignment throughout the distributional spectrum provides compelling visual evidence that the residuals conform closely to the expected normal distribution, thereby satisfying one of the critical assumptions underlying valid statistical inference in regression analysis.

The central portion of the distribution, which carries the greatest weight in parameter estimation and hypothesis testing, demonstrates particularly strong alignment with the

Table 8: Regression Model Summary.

Model	R	R <sup>2</sup>	Adjusted R <sup>2</sup>	Std. Error of Estimate	Durbin-Watson
1	0.849a	0.721	0.718	3.674	1.98

a. Predictors: (Constant), F4, F1, F2, F3

Table 9: ANOVA for the Regression Model.

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	15520.34	4	3880.09	287.41	0.000b
Residual	5994.66	444	13.50		
Total	21515.00	448			

b. Predictors: (Constant), F4, F1, F2, F3

Table 10: Regression Model Coefficients.

Model	Unstandardized Coef.	Standardized Coef.	t	Sig.
	B	Std. Error	Beta	
(Constant)	42.310	0.173		244.57 0.000
F1: Interactive Engagement	2.481	0.174	0.358	14.26 0.000
F2: Collaborative Learning	2.437	0.174	0.352	14.01 0.000
F3: Asynchronous Comm.	2.915	0.174	0.421	16.75 0.000
F4: Personalized Learning	1.315	0.174	0.19	7.56 0.000

Dependent Variable: Total Critical Thinking Score.

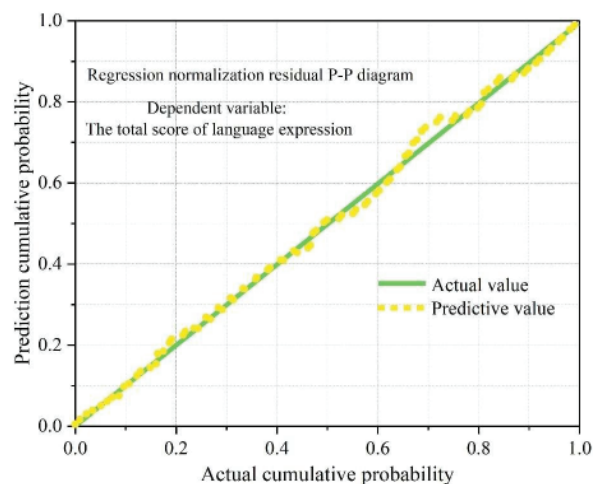


Figure 5: Normal P-P Plot of Regression Standardized Residuals.



diagonal reference line. For actual cumulative probabilities ranging from 0.22 through 0.30, the corresponding predicted values of 0.66, 0.72, 0.78, 0.84, and 0.90 exhibit minimal deviation from the expected trajectory. This concentration of well-behaved residuals in the region where the majority of observations reside is especially important, as it ensures that the core of the distribution conforms to normality assumptions even if minor departures occur at the extremes. The absence of systematic curvature or S-shaped patterns in this central region indicates that the error term distribution possesses the necessary properties for unbiased parameter estimation and accurate confidence interval construction.

Examination of the upper tail of the distribution further reinforces the conclusion of approximate normality. The coordinates spanning actual probabilities from 0.32 to 0.34 demonstrate predicted values of 0.96 and 1.00, respectively, maintaining the pattern of close alignment with the diagonal. The absence of substantial divergence at either extreme tail is particularly noteworthy, as residual distributions often exhibit departures from normality at the boundaries due to floor or ceiling effects in the dependent variable. The well-behaved nature of both tails suggests that the regression model performs consistently across the entire spectrum of predicted values, with no evidence of systematic underprediction or overprediction that might indicate violation of the normality assumption or presence of unmodeled nonlinearities.

The confirmation of residual normality through the P-P plot analysis carries substantive implications for the validity and reliability of the regression findings presented in this study. Satisfying the normality assumption ensures that the standard errors of the regression coefficients are accurately estimated, that hypothesis tests (t-tests and F-tests) maintain their nominal Type I error rates, and that confidence intervals constructed around parameter estimates possess the stated coverage probabilities. Furthermore, the normality of residuals provides indirect confirmation that the linear functional form specified in the regression model appropriately captures the relationship between the predictor variables and the dependent variable, as systematic departures from normality often signal model misspecification. Collectively, the evidence from Figure 4, in conjunction with the previously reported skewness and kurtosis statistics, provides robust justification for the statistical inferences drawn from the regression analysis and reinforces confidence in the substantive conclusions regarding the relationship between university language education and students' expressive abilities.

## Conclusion

This paper provides a robust quantitative analysis of the relationship between digital pedagogy and critical thinking skills in higher education. Using a structured questionnaire and employing Principal Component Analysis (PCA) to distill key dimensions of digital practice, we constructed a multivariate linear regression model to assess their impact on students' critical thinking. The results lead to several important conclusions:

1. The PCA effectively reduced the complexity of digital pedagogy into four meaningful and uncorrelated principal components: "Interactive Engagement," "Collaborative Learning," "Asynchronous Communication," and "Personalized Learning." These components, explaining a substantial 82.47% of the total variance, provide a clear framework for understanding different facets of technology use in education.
2. The mean critical thinking score (42.31 out of 60) indicates a moderate level of proficiency among students. However, the significant range of scores (37 points) and standard deviation (6.92) reveal considerable disparities in students' higher-order cognitive skills, underscoring the need for targeted pedagogical interventions.
3. The multivariate regression analysis revealed a powerful and significant relationship. The four digital pedagogy components together accounted for an impressive 72.1% of the variance in critical thinking scores ( $R^2 = 0.721$ ). All four components were significant positive predictors, confirming that a richer and more diverse engagement with digital learning environments is strongly associated with enhanced critical thinking. While this  $R^2$  value is notably high for survey-based educational research, it is important to interpret this finding with appropriate caution; the explained variance reflects the strength of the statistical association between the measured constructs within this specific sample, and given the cross-sectional design and reliance on self-reported data, these results should not be interpreted as evidence of causal relationships.
4. Notably, "Asynchronous Communication" (e.g., discussion forums, reflective blogs) emerged as the strongest predictor, followed by "Collaborative Learning" (e.g., online group projects). This highlights the particular value of technologies that promote reflection, dialogue, and social knowledge construction in fostering analytical and evaluative skills.

This study provides strong empirical evidence that the pedagogical application of digital tools, especially those promoting collaboration and reflection, is directly and positively linked to the development of critical thinking in university students. The derived regression model offers a valuable predictive tool for educators and curriculum designers seeking to optimize digital learning strategies to cultivate these essential competencies. The findings advocate for a shift from simply using technology for content delivery to leveraging it as a medium for active, social, and analytical learning.

## Recommendations

Based on the findings of this study, the following recommendations are proposed for educational practitioners, institutional leaders, and future researchers. Each recommendation has been refined to include concrete, actionable examples that can be directly implemented in educational practice.

## For educators and instructional designers

### 1. Prioritize asynchronous discussion forums in course design:

**Actionable Example:** Implement a structured "Forum Friday" activity where students must post an initial analysis of a weekly reading by Thursday, then provide evidence-based critiques to at least two peers by Sunday. Require students to support their critiques with specific references to course materials and to identify strengths and weaknesses in their peers' arguments using a provided rubric focused on logical reasoning and evidence quality.

**Additional Application:** Design a "Perspective Taking" discussion thread where students are assigned different stakeholder positions on a controversial topic (e.g., environmental policy, ethical AI use) and must construct arguments from that perspective while responding to opposing viewpoints.

### 2. Design collaborative online projects with authentic problem-solving tasks:

**Actionable Example:** Create a semester-long "Consultancy Project" where student teams (4-5 members) work with a real or simulated community partner to solve an authentic problem. Implement the "Jigsaw Method" wherein each team member becomes an expert on a different aspect of the problem (e.g., research, data analysis, stakeholder impact, feasibility) and is responsible for teaching their component to the team. Include individual accountability through weekly reflection logs documenting each member's contribution and analytical process.

**Additional Application:** Design a "Collaborative Critique" assignment where teams exchange draft projects mid-semester and provide structured peer feedback using a critical evaluation framework that requires them to identify assumptions, evaluate evidence quality, and suggest alternative approaches.

### 3. Integrate interactive multimedia engagement strategically

**Actionable Example:** Move beyond passive video watching by embedding interactive "pause and predict" points within lecture videos using tools like PlayPosit or H5P. At critical junctures, require students to make predictions about experiment outcomes, answer analytical questions about causal mechanisms, or identify flaws in presented arguments before proceeding.

**Additional Application:** Use PhET interactive simulations in science courses where students must manipulate variables, document their hypotheses, record outcomes, and explain discrepancies between predicted and actual results in a structured lab notebook template that emphasizes analytical reasoning.

### 4. Implement adaptive learning technologies with deliberate scaffolding

**Actionable Example:** When using adaptive platforms like ALEKS or Smart Sparrow, design custom "reflection checkpoints" that activate after students complete each mastery unit. At these checkpoints, require students to respond to prompts such as: "Explain the most common error students make in this type of problem and why it occurs," or "Create your own practice problem that tests the same concept and provide a step-by-step solution with reasoning."

**Additional Application:** Program the adaptive system to present students with "what-if" scenarios when they answer incorrectly, prompting them to trace their reasoning steps and identify where their thinking diverged from the correct approach before proceeding to remediation content.

### 5. Develop explicit critical thinking learning objectives for digitally-mediated activities

**Actionable Example:** For each digital learning activity, provide students with a "Critical Thinking Focus" box that specifies the targeted skill. For example, in a discussion forum assignment, state: "This activity targets Evaluation skills. You will practice assessing the credibility of sources and judging the logical consistency of arguments. Your posts will be evaluated using the attached rubric, focusing on evidence quality and reasoning clarity."

**Additional Application:** Create a shared "Critical Thinking Skills Matrix" that maps each course activity to specific sub-skills (Analysis, Evaluation, Inference, Explanation) and shares this matrix with students at the beginning of the semester, allowing them to track their own development across these dimensions.

## For institutional leaders and policymakers

### 1. Invest in faculty development programs targeting higher-order cognitive outcomes

**Actionable Example:** Establish a semester-long "Digital Pedagogy for Critical Thinking" faculty learning community where participants redesign one of their courses using the four-component framework from this study. Provide each participant with an instructional designer to help implement structured discussion forums, collaborative projects, and adaptive learning checkpoints. Require participants to share pre- and post-implementation student work samples demonstrating critical thinking growth.

**Additional Application:** Create a "Teaching Innovation Grant" program that funds faculty projects explicitly designed to enhance critical thinking through digital pedagogy, with priority given to proposals



that incorporate asynchronous communication and collaborative learning components.

## 2. Establish institutional frameworks for assessing critical thinking

**Actionable Example:** Implement a longitudinal assessment program using the Valid Assessment of Learning in Undergraduate Education (VALUE) Critical Thinking Rubric, applied to student work samples collected at entry (first-year), midpoint (sophomore/junior), and exit (senior) stages. Use a standardized submission process where random samples of student work from designated "critical thinking intensive" courses are anonymously scored by trained faculty raters.

**Additional Application:** Develop digital portfolios where students curate work demonstrating their critical thinking development, accompanied by reflective essays explaining how specific digital learning activities contributed to their growth in analysis, evaluation, inference, and explanation skills.

## 3. Allocate resources toward integrated learning platforms

**Actionable Example:** Invest in learning management system enhancements that integrate discussion analytics (e.g., quality of posts, depth of threads, patterns of student interaction) with early alert systems. Provide faculty with automated dashboards showing which students are demonstrating superficial engagement (e.g., surface-level posts, limited response to peers) and suggesting targeted interventions such as personalized feedback or structured discussion prompts.

**Additional Application:** Pilot an "AI Teaching Assistant" tool that monitors student contributions in discussion forums and provides real-time prompts such as: "Consider providing evidence to support your claim" or "You've stated your position—now try addressing a potential counterargument."

## 4. Develop institutional policies addressing AI integration

**Actionable Example:** Replace blanket prohibitions on generative AI with a "scaffolded AI use" policy that specifies when and how AI tools may be used. For example, policy might state: "AI may be used for brainstorming initial ideas and checking grammar, but all analytical writing must be student-generated. Students must submit a 'process statement' documenting how AI was used and reflecting on how their own thinking evolved through the writing process."

**Additional Application:** Create institutional guidelines requiring that any course using AI tools must include explicit instruction on "critical AI literacy"—teaching students to evaluate AI-generated outputs for accuracy, bias, and logical consistency, and to identify instances where AI overconfidence might mask analytical gaps.

## For future researchers

### 1. Conduct longitudinal studies tracking critical thinking development

**Actionable Example:** Design a multi-year study tracking a cohort of students from first-year entry through graduation, collecting critical thinking assessments at four time points (entry, end of year 1, end of year 2, exit) and linking these to learning analytics data capturing engagement with specific digital pedagogical components across courses.

### 2. Employ experimental and quasi-experimental designs

**Actionable Example:** Implement a randomized controlled trial in a large introductory course where sections are randomly assigned to either a "structured discussion forum" condition (with required peer critique protocols) or a "traditional discussion" condition (optional, unmoderated forums), comparing pre-post critical thinking gains while controlling for prior ability and motivation.

### 3. Investigate disciplinary variations

**Actionable Example:** Conduct a comparative study examining how asynchronous communication influences critical thinking development in humanities courses (where interpretation and argumentation are central) versus STEM courses (where problem-solving and evidence evaluation are emphasized), using discipline-specific critical thinking assessments.

### 4. Examine mediating and moderating variables

**Actionable Example:** Develop and test a structural equation model examining whether metacognitive awareness mediates the relationship between asynchronous discussion engagement and critical thinking gains, using validated instruments to measure metacognition, self-regulated learning, and digital literacy as potential mediators.

### 5. Validate and refine measurement instruments

**Actionable Example:** Conduct a cross-national validation study of the Digital Pedagogy Engagement Scale used in this research, administering the instrument to student samples in at least five countries across different continents and using multi-group confirmatory factor analysis to establish measurement invariance.

### 6. Investigate the differential effects of generative AI tools

**Actionable Example:** Design an experimental study comparing three conditions in a writing-intensive course: (1) unconstrained AI access, (2) scaffolded AI access with structured reflection requirements, and (3) no AI access. Compare pre-post critical thinking gains and analyze student writing samples for evidence of analytical depth, original thinking, and metacognitive reflection.



## Acknowledgement

The authors wish to express their sincere gratitude to the students and faculty who participated in this study, generously sharing their time and insights. Special appreciation is extended to the academic staff at the participating universities for facilitating data collection and providing access to student populations. The authors acknowledge the valuable contributions of research assistants who assisted with questionnaire administration, data entry, and preliminary analyses. Portions of this research were presented at the International Conference on Digital Pedagogy (ICDP 2026) in Indonesia, and the authors appreciate the insightful comments received from conference participants that strengthened the final manuscript.

## Conflict of interest

The authors declare no actual or potential conflict of interest concerning the research, authorship, and/or publication of this article. The research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. All authors have reviewed and approved the final version of the manuscript and agree with its submission for publication.

## References

1. Organisation for Economic Co-operation and Development (OECD). Digital Education Outlook 2026: AI, Cognition, and the Future of Learning. Paris: OECD Publishing; 2026. Available from: [https://www.oecd.org/en/publications/oecd-digital-education-outlook-2026\\_062a7394-en.html](https://www.oecd.org/en/publications/oecd-digital-education-outlook-2026_062a7394-en.html)
2. Molenaar I, Kirschner PA. The mirage of false mastery: How generative AI diminishes metacognitive engagement in higher education. *Comput Educ Artif Intell.* 2025;6(1):100234. Available from: <https://www.sciencedirect.com/journal/computers-and-education-artificial-intelligence>
3. World Economic Forum. The Future of Jobs Report 2026: Skills for a Digital Era. Geneva: WEF Publishing; 2026.
4. Arum R, Roksa J. *Academically Adrift at Twenty: A Reassessment of Critical Thinking Gains in Higher Education.* Chicago: University of Chicago Press; 2024.
5. Paul R, Elder L. *Critical Thinking: Tools for Taking Charge of Your Professional and Academic Life.* 4th ed. Tomales (CA): Foundation for Critical Thinking Press; 2025. Available from: <https://www.criticalthinking.org/store/products/critical-thinking-tools-for-taking-charge-of-your-learning-amp-your-life-2nd-edition/143>
6. Abrami PC, Bernard RM, Borokhovski E, Waddington DI, Wade CA, Persson T. Instructional interventions affecting critical thinking skills and dispositions: A staged meta-analysis update 2015-2025. *Rev Educ Res.* 2025;95(2):215-268. Available from: <https://journals.sagepub.com/home/rev>
7. Luckin R, Cukurova M. Intelligence-driven education: Integrating AI and learning analytics for personalized critical thinking development. *Int J Artif Intell Educ.* 2026;36(1):45-72. Available from: <https://link.springer.com/journal/40593>
8. Facione PA, Gittens CA. *Think Critically: A Framework for Reasoning in the Age of AI.* 5th ed. London: Pearson Education; 2025. Available from: <https://www.pearson.com/en-us/subject-catalog/p/think-critically/>
9. Schober P, Vetter TR. Linear regression in medical research. *Anesth Analg.* 2021;132(1):108-109. Available from: [https://journals.lww.com/anesthesia-analgesia/fulltext/2021/01000/linear\\_regression\\_in\\_medical\\_research.18.aspx](https://journals.lww.com/anesthesia-analgesia/fulltext/2021/01000/linear_regression_in_medical_research.18.aspx)
10. Wen J, Wei X, He T, Zhang S. Regression analysis on the influencing factors of the acceptance of online education platform among college students. *Ing Syst Inf.* 2020;25(5). Available from: <https://www.iieta.org/journals/isi/paper/10.18280/isi.250506>
11. Shavit Y, Edelman B, Axelrod B. Causal strategic linear regression. In: *International Conference on Machine Learning; 2020 Nov.* PMLR; 2020;8676-8686. Available from: <http://proceedings.mlr.press/v119/shavit20a.html>
12. Lumley T, Scott A. Fitting regression models to survey data. *Stat Sci.* 2017;32(2):265-278. Available from: <https://projecteuclid.org/journals/statistical-science/volume-32/issue-2/Fitting-Regression-Models-to-Survey-Data/10.1214/16-STS605.full>
13. Baždarić K, Šverko D, Salarić I, Martinović A, Lucijanić M. The ABC of linear regression analysis: What every author and editor should know. *Eur Sci Ed.* 2021;47. Available from: <https://ese.arphahub.com/article/63780/>
14. Liu Y, Zhang S. Fast quantum algorithms for least squares regression and statistical leverage scores. *Theor Comput Sci.* 2017;657:38-47. Available from: <https://doi.org/10.1016/j.tcs.2016.05.044>
15. Bian W, Dong W, Zheng Q, Gu Q, Bian S, Yang Y. Fast weighted least squares for detail and tone enhancement of medical images. *Digit Health.* 2024. Available from: <https://doi.org/10.1177/20552076241306272>
16. Li H, Dai X, Zhou L, Wu Q, Deveci M, Pamucar D. A least-squares framework for developing interval type-2 fuzzy semantics. *Appl Soft Comput.* 2024;112293. Available from: <https://doi.org/10.1016/j.asoc.2024.112293>
17. Gong Y, Lyu B, Gao X. Research on teaching: a bibliometric analysis. *Asia Pac Educ Res.* 2018;27(4):277-289. Available from: <https://doi.org/10.1007/s40299-018-0385-2>
18. Xie Y, Ryder L, Chen Y. Using interactive virtual reality tools in an advanced Chinese language class: A case study. *TechTrends.* 2019;63:251-259. Available from: <https://doi.org/10.1007/s11528-019-00389-z>
19. Xu Q, Peng H. Investigating mobile-assisted oral feedback in teaching Chinese as a second language. *Comput Assist Lang Learn.* 2017;30(3-4):173-182. Available from: <https://doi.org/10.1080/09588221.2017.1297836>
20. Shabbir MS, Wisdom O. The relationship between corporate social responsibility, environmental investments, and financial performance. *Environ Sci Pollut Res.* 2020;27(32):39946-39957. Available from: <https://doi.org/10.1007/s11356-020-10217-0>
21. Zhao Q. Modeling and analysis method of national fitness big data for basketball projects based on a multivariate statistical model. *Secur Commun Netw.* 2022. Available from: <https://doi.org/10.1155/2022/2591633>
22. Facione PA. *Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction (The Delphi Report).* California: The California Academic Press; 1990. Available from: [https://www.researchgate.net/publication/242279575\\_Critical\\_Thinking\\_A\\_Statement\\_of\\_Expert\\_Consensus\\_for\\_Purposes\\_of\\_Educational\\_Assessment\\_and\\_Instruction](https://www.researchgate.net/publication/242279575_Critical_Thinking_A_Statement_of_Expert_Consensus_for_Purposes_of_Educational_Assessment_and_Instruction)
23. Wang J, Zhang C, Zhao W, Huang X, Nie F. Fast anchor graph optimized projections with principal component analysis and entropy regularization. *Inf Sci.* 2025:121797. Available from: <https://doi.org/10.1016/j.ins.2024.121797>
24. Stanley TD, Doucouliagos H, Steel P. Does ICT generate economic growth? A meta-regression analysis. *J Econ Surv.* 2018;32(3):705-726. Available from: <https://doi.org/10.1111/joes.12211>
25. Yamaguchi A, Arai K, Aisnada ANE, Lee JE, Kitadai N, Nakamura R, Miyauchi M. Multiregression analysis of CO<sub>2</sub> electroreduction activities on metal sulfides. *J Phys Chem C.* 2022;126(5):2772-2779. Available from: <https://doi.org/10.1021/acs.jpcc.1c08993>
26. Shastry A, Sanjay HA, Bhanusree E. Prediction of crop yield using regression techniques. *Int J Soft Comput.* 2017;12(2):96-102. Available from: <https://doi.org/ijscmp.2017.96.102>