International Journal of Learning, Teaching and Educational Research
Vol. 24, No. 12, pp. 118-134, December 2025
https://doi.org/10.26803/ijlter.24.12.6
Received *Sept 24, 2025; Revised Oct 26, 2025; Accepted Oct 29, 2025*

Between Assistance and Dependence: Artificial Intelligence and Critical Thinking. A Study on First-Year University Students

Elvis Amado Vereau Amaya, Annie Rosa Soto Zavaleta, Irene Gregoria Vásquez Luján
Universidad Nacional del Santa
Av. Universitaria S/N, Nuevo Chimbote, Perú

Elmer González López
Universidad Nacional de Trujillo
Av. Juan Pablo II, Trujillo, Perú

Janeth Tomanguilla Reyna
Universidad César Vallejo
Av. Larco 1770, Trujillo, Perú

Jorge David Ríos Gonzales*
Universidad Nacional de Trujillo
Av. Juan Pablo II, Trujillo, Perú

Abstract. This study examined the relationship between dependence on generative AI tools and the dimensions of critical thinking among first-semester students at a university in Trujillo, Peru. A quantitative, correlational, cross-sectional design was applied to a sample of 200 students, using validated Likert-scale instruments for AI dependence and for interpretation, explanation, inference, analysis, and evaluation. The results show moderate levels of AI use, primarily as support, alongside an uneven critical-thinking profile. AI dependence is negatively associated with the overall critical-thinking score and more strongly with inference and evaluation, whereas interpretation and explanation remain comparatively stronger. The regression models explain a meaningful share of the variance and suggest that unmediated AI use may displace self-regulatory processes in novice students. The study contributes empirical evidence from the Peruvian context—underrepresented in the regional literature—and offers practical guidelines for integrating AI as a verifiable pedagogical scaffolding that requires reconstructing reasoning and verifying sources. Curricular interventions aiseacher training in

*Corresponding author: Jorge David Ríos Gonzales; *jriosg@unitru.edu.pe*

critical digital literacy are recommended, as are longitudinal and experimental studies to strengthen causal inference.

Keywords: ChatGPT; TPACK; artificial intelligence; digital pedagogy; higher education

1. Introduction

The rapid diffusion of generative AI in higher education has heightened the core tension between technological assistance and cognitive dependence. This tension is particularly acute for first-year students whose still-forming academic habits can be displaced by automated outputs, with the potential erosion of critical thinking and self-regulation (Adiguzel et al., 2023; Rasul et al., 2023). While tools such as ChatGPT can streamline tasks and support learning, unmediated reliance raises concerns about intellectual autonomy and academic integrity, precisely the capacities universities are mandated to cultivate.

Despite growing international scholarship, Latin America remains underrepresented in evidence on AI and student learning. In Peru, structural constraints—including persistent digital divides, uneven professional training, and fragmented institutional policies—complicate meaningful classroom integration, especially outside major metropolitan centers (Atchley et al., 2024; OECD, 2024; Yue et al., 2024). Grounding the study in Trujillo therefore provides localized empirical insight that can inform ethical and pedagogically sound adoption in comparable regional settings.

Previous work often highlights creativity, efficiency, and productivity gains among students, typically in advanced systems, but offers limited empirical analysis of how AI reliance relates to critical-thinking dimensions—interpretation, analysis, inference, evaluation, and explanation—at the outset of university, and virtually none in Latin American contexts (Flores-Vivar & García-Peñalvo, 2023; Lo, 2023). Moreover, much of the literature remains descriptive, seldom testing associations between overreliance on machine-generated content and the erosion of higher-order cognitive skills (Kasneci et al., 2023; Kooli, 2023). This gap is consequential for first-year cohorts, who may be most vulnerable to substituting reflective inquiry with automated outputs (Alnasib, 2023; Dergaa et al., 2023).

Against this backdrop, the study examines the relationship between students' reliance on generative AI and critical thinking among first-year undergraduates in Trujillo, Peru. The objectives are threefold: to identify the prevalence and usage patterns of generative AI tools among first-year students; to evaluate levels of critical thinking across analysis, interpretation, inference, evaluation, and explanation; and to determine the direction and magnitude of the association between AI reliance and critical thinking, both overall and by dimension (Almulla & Ali, 2024; Yue et al., 2024; Alrishan, 2023). As such, the inquiry clarifies the empirical gap, centers the vulnerability of first-year students as the primary rationale for the population choice, and positions the Peruvian case to contribute regionally relevant evidence to international debates.

2. Literature Review

2.1 Artificial Intelligence in Higher Education

The integration of artificial intelligence in higher education has accelerated with the emergence of generative tools such as ChatGPT, Bard, and Copilot, which have introduced a broadened repertoire of supports for students and faculty—from text generation to content synthesis and the simulation of academic dialogues. Recent literature documents that these systems can improve efficiency, personalize learning, and foster creativity in instructional design, relieving routine workloads and opening space for higher-value cognitive activities (Banihashem et al., 2024; Lo, 2023). Nevertheless, alongside these promises come legitimate concerns about plagiarism, academic integrity, and the risk of shallow learning when dependence on automated outputs becomes normalized without pedagogical mediation, especially in contexts where assessment privileges products over processes (Bitzenbauer, 2023).

In Latin America, adoption shows additional nuances. In analyzing ChatGPT integration through the TPACK framework, Ríos Gonzales et al. (2025) show that faculty demographic profiles and levels of techno-pedagogical knowledge significantly condition both the selection of AI-based activities and the quality control of generated products. This evidence suggests that adoption challenges are not exclusively technical but also pedagogical and institutional, which makes it necessary to understand how students engage with these tools and what the cognitive implications of sustained use are in settings marked by infrastructure and training gaps.

2.2 Critical Thinking in University Students

Critical thinking is recognized as a cardinal competency in university education and a pillar for responsible professional insertion. Drawing on the Delphi model, Facione (2015) defines it as an integrated set of skills—analysis, interpretation, inference, evaluation, and explanation—that underpin reflective judgment and informed decision-making. In information-saturated societies and evidence-based production environments, these skills are indispensable for discriminating sources, articulating arguments, and sustaining justified positions (Bukar et al., 2024).

Even so, first-year students often exhibit only incipient levels of critical thinking, due to academic habits still in consolidation and not-yet-stabilized metacognitive strategies, all of which hinders sustained deep reading, argumentation, and verification (Castagnola Rossini et al., 2025). When generative tools provide immediate answers without an instructional frame that requires reconstructing the underlying reasoning, then incentives for reflective inquiry weaken. Hence, strengthening critical thinking in the early stages remains a curricular priority and a research problem that calls for designs and metrics tailored to university entry levels.

2.3 AI Dependence and Cognitive Skills

Beyond access, the literature warns about the shift from scaffolding to substitution when AI use becomes routine and unmediated. Overreliance on automated content tends to reduce student participation in elaborative activities while compromising higher-order skills—particularly inference and evaluation—which are sensitive to information quality and the strength of justifications (Celik, 2023). Chamorro-Atalaya et al. (2023) argue that although AI can operate as cognitive support, in the absence of clear formative goals and instructor feedback it can displace key processes of self-regulation and autonomous problem-solving. In the same vein, De Jesus et al. (2024) warns that digital overdependence promotes passivity and reduces students' intellectual agency.

Latin American evidence adds a relevant nuance: when AI is embedded in structured didactic sequences with tasks for verification and reasoning reconstruction, it can enhance the research skills tied to critical thinking—information analysis, methodological reasoning, and data interpretation—as shown in an intervention in Trujillo, Peru (Ríos Gonzales et al., 2025). In short, the effect is not uniform; it depends on pedagogical scaffolding, evaluative expectations, and learners' stage of development—variables that often differ across institutions and regions.

2.4 Latin American Perspectives

In Latin America, research on AI in higher education is advancing, but is doing so with thematic dispersion and methodological heterogeneity. Persistent digital divides, asymmetries in infrastructure, and lags in faculty training condition the possibility of integrating AI practices meaningfully and sustainably (Duong, 2024). The absence—or incipience—of clear institutional policies also generates variability across universities and programs, with disparate instructional strategies and weakly standardized assessment criteria (Essien et al., 2024).

Although Peruvian studies provide evidence on faculty adoption and student skill development (Ampo et al.,2025; Ríos Gonzales et al., 2024), there are still few studies that empirically link AI dependence to specific dimensions of critical thinking across diverse institutional contexts. Compared with North America, Europe, and parts of Asia—where technological infrastructure and faculty professional development are more consolidated—Latin American institutions operate under constraints that affect not only access but also quality of use and pedagogical alignment, thereby modulating whether AI functions as scaffolding or devolves into a cognitive crutch.

2.5 Research Gap and Study Contribution

Taken together, while international evidence documents opportunities and risks, decisive gaps still persist. On the one hand, there is limited empirical analysis of how AI dependence relates to concrete dimensions of critical thinking at the outset of university study, when these processes are most malleable. On the other, the underrepresentation of Latin America makes it difficult to understand how institutional and pedagogical constraints modify these relationships. To address these gaps, the present study examines the association between dependent use of generative AI tools and critical thinking among first-year students in Trujillo, Peru, with the purpose of providing localized evidence that specifies the conditions under which AI acts as formative support and those in which it tends to substitute essential cognitive processes, while also informing curricular and university policy decision.

2.6 Theoretical Framework

The study is grounded in complementary perspectives that delimit the constructs and specify how they are measured. Following Facione's (2015) Delphi model, critical thinking is operationalized as five observable skill domains—analysis, interpretation, inference, evaluation, and explanation—captured through subscale scores that reflect students' agreement with behaviors indicative of each domain (e.g., weighing evidence, drawing warranted conclusions, justifying claims). In line with Bloom's revised taxonomy (Aldawsari et al., 2023), item phrasing targets upper-level processes (analyze/evaluate/create) rather than recall or comprehension, so that the composite and dimension scores represent higher-order performance rather than lower-order achievement. Within this framework, the dependent variable is the total critical-thinking score and its five-dimension scores, treated as continuous indices derived from validated Likert-type items.

The independent variable, AI reliance, is defined—drawing on Jose et al. (2025), cognitive offloading and Vygotskian scaffolding—as the extent to which students report using generative AI to substitute (rather than support) core cognitive operations. Accordingly, the AI-reliance index aggregates frequency and purpose-of-use indicators (e.g., drafting vs. editing; generating answers vs. verifying sources), plus two practice markers aligned with self-regulated learning (Zimmerman, 2002): planning/monitoring before AI use and post-use verification. Higher scores therefore indicate greater substitution and reduced self-regulatory engagement, consistent with risks highlighted by Cohen, L., Manion & Morrison (2018).

This yields a theoretically anchored gradient from scaffolding (guided support with planning and verification) to substitution (automation without reconstruction), which we test empirically against the critical-thinking outcomes. Finally, the surface-versus-deep learning lens (Anand, 2024) informs interpretation: patterns of high substitution and low verification are expected to align with lower inference and evaluation scores, whereas support-oriented use with monitoring/verification should be neutral or positively associated with higher-order outcomes (Funda & Mbangeleli, 2024; Goh & Sandars, 2024; Gouia-Zarrad & Gunn, 2024; Fuchs, 2023).

3. Methodology

3.1 Research Design

This study employed a quantitative, correlational, cross-sectional design to examine the association between students' reliance on artificial intelligence (AI) tools and their critical-thinking skills. A correlational approach was deemed appropriate because the goal was to estimate the strength and direction of relationships rather than establish causality (Creswell & Creswell, 2023). The cross-sectional design enabled data collection at a single time point, capturing current patterns of AI use and contemporaneous critical-thinking levels.

3.2 Population and Sample

Participants were 200 first-year undergraduates enrolled in general education courses at a private university in Trujillo, Peru, during the first semester of 2025. First-year students were targeted because emerging academic habits may increase vulnerability to cognitive substitution when using generative AI. A non-probability purposive sampling strategy was used; inclusion criteria required active enrollment in the first academic cycle and prior exposure to at least one generative AI tool. However, while this strategy is theoretically justified, it limits external validity—in particular, it may introduce bias toward students in private higher education in Trujillo—so generalization to public universities or other regions should be made with caution.

3.3 Variables and Operationalization

Two constructs were analyzed. AI dependence was defined as the extent to which students substitute core cognitive operations with generative AI (e.g., drafting answers rather than planning, monitoring, or verifying). It was operationalized via frequency of use, purpose of use (support vs. substitution), and self-perceived reliance. Critical thinking followed Facione's Delphi model (2015) and comprised analysis, interpretation, inference, evaluation, and explanation. Both constructs were measured with validated Likert-type instruments adapted to the Peruvian higher-education context to preserve conceptual alignment and cultural/curricular relevance.

3.4 Instruments

Data were collected with two standardized questionnaires. The AI Dependence Scale, adapted from Kasneci et al. (2023) and Rasul et al. (2023), assessed frequency, intensity, purpose of use, and perceived reliance on generative AI. The Critical Thinking Skills Test, based on Facione (2015), assessed the five dimensions noted above. Content validity was established by expert review, and a pilot with 30 first-year students confirmed clarity and internal consistency. Reliability evidence included Cronbach's alpha above conventional thresholds for both instruments ($\alpha = .88$ for AI dependence; $\alpha = .91$ for critical thinking).

Beyond internal consistency, we conducted factor-analytic validation. In the pilot, an exploratory factor analysis (EFA) indicated adequate sampling (KMO = .86) and significant sphericity (Bartlett's $\chi^2(190) = 1425.7$, $p < .001$), with items loading $\geq .50$ on their intended dimensions and minimal cross-loadings. In the main sample, confirmatory factor analysis (CFA) showed satisfactory fit: CFI = .95, TLI = .94, RMSEA = .05 (90% CI .04–.07), SRMR = .04. Composite reliability (CR) ranged from .78 to .88 across subscales, and average variance extracted (AVE) ranged from .52 to .62, supporting convergent and discriminant validity. These results align the instruments with the intended theoretical structure.

3.5 Data Collection Procedure

Following Institutional Review Board approval, students received information about aims, risks/benefits, anonymity, and voluntariness, and provided written informed consent. Questionnaires were administered during scheduled class sessions in paper and digital formats to maximize coverage while preserving confidentiality. Participation was voluntary, and withdrawal was permitted at any time without penalty.

3.6 Data Analysis

Analyses were conducted in SPSS v27. Descriptive statistics (means, standard deviations, frequencies) summarized AI-use patterns and critical-thinking levels. Although single Likert items are ordinal, composite scores based on multiple items can be treated as approximately interval when internal consistency is high and sample size is adequate (Han, 2024; Iqbal & Rahman, 2024). Associations between AI dependence and critical-thinking outcomes were estimated with Pearson's r; as a robustness check, Spearman's ρ was inspected when normality assumptions were not met.

To examine predictive effects, we ran multiple linear regressions with overall and dimension-level critical thinking as outcomes and AI dependence as the focal predictor. We checked model assumptions (linearity, normality of residuals, homoscedasticity), assessed multicollinearity (VIF < 3), and reported effect sizes and 95% confidence intervals alongside p-values ($\alpha = .05$). Missing data were minimal (< 3%) and handled via listwise deletion under plausible MCAR; sensitivity analyses yielded substantively identical results.

4. Results

4.1 Descriptive Statistics

The sample comprised 200 first-year students (52% female, 48% male; M age = 18.9, SD = 1.1). AI use was widespread yet mostly moderate: 78% reported at least weekly use and 12% daily use. Most uses were support-oriented (e.g., summarizing, clarifying concepts), while 30% reported substitution to complete assignments (see Table 1 for usage distribution and Figure 1 for frequency patterns).

Table 1: Patterns of AI usage among participants

AI Use Frequency	% of Students	Main Purpose of Use
Rarely (monthly)	10%	Curiosity, general questions
Weekly	56%	Academic support
Several times/week	22%	Writing support, summaries
Daily	12%	Completing assignments

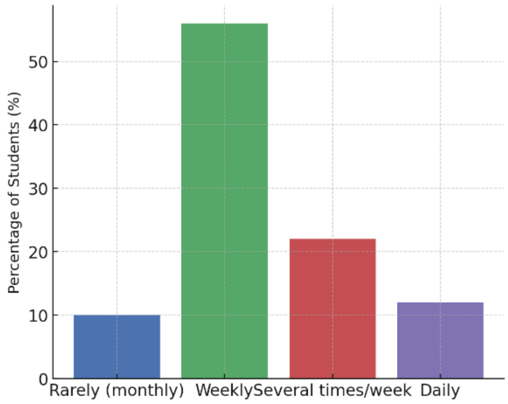

Figure 1: Frequency of AI Use among Students (N = 200)

Descriptive statistics for the study variables are summarized in Table 2. AI dependence was moderate (M = 3.12, SD = 0.64), while critical thinking showed an uneven profile (M = 3.45, SD = 0.58), with interpretation comparatively higher (M = 3.61, SD = 0.62) and inference lower (M = 3.28, SD = 0.66), anticipating the dimension-level patterns explored in subsequent analyses.

Table 2: Descriptive statistics for study variables

Variable/Dimension	Mean	SD
AI Dependence	3.12	0.64
Critical Thinking (total)	3.45	0.58
– Analysis	3.42	0.61
– Interpretation	3.61	0.62
– Inference	3.28	0.66
– Evaluation	3.49	0.57
– Explanation	3.49	0.59

4.2 Reliability of the Instruments

Internal consistency was high (see Table 3): AI Dependence α = .88; Critical Thinking (total) α = .91; subscales α = .82–.86. Factor-analytic evidence supported construct validity: the pilot EFA showed adequate sampling (KMO = .86) and significant sphericity (Bartlett's $\chi^2(190) = 1425.7$, $p < .001$) with item loadings ≥ .50 on intended factors and minimal cross-loadings; in the main sample, the CFA indicated good fit (CFI = .95, TLI = .94, RMSEA = .05 [90% CI .04–.07], SRMR = .04). Composite reliability (CR = .78–.88) and average variance extracted (AVE = .52–.62) supported convergent and discriminant validity across all subscales.

Table 3: Reliability coefficients of study instruments

Instrument / Dimension	Cronbach's α	CR	AVE	Factor loadings
AI Dependence Scale (total)	.83	.86	.55	.58-.79
Critical Thinking Test (total)	.91	.89	.57	.60-.82
– Analysis	.83	.82	.52	.57-.73
– Interpretation	.84	.83	.54	.59-.76
– Inference	.86	.88	.62	.65-.82
– Evaluation	.82	.80	.53	.56-.74
– Explanation	.84	.81	.52	.55-.72

4.3 Levels of Critical Thinking

Performance across the five dimensions was uneven. Interpretation and Explanation showed the highest means, whereas Inference was lowest. This pattern was more pronounced among students with higher-frequency and substitution-oriented AI use, suggesting displacement of inferential and evaluative processes (see Table 4 for level distributions and Figure 2 for the mean profile).

Table 4: Levels of critical thinking by dimension

Dimension	Low (%)	Medium (%)	High (%)
Analysis	18	55	27
Interpretation	12	52	36
Inference	25	58	17
Evaluation	20	54	26
Explanation	15	53	32

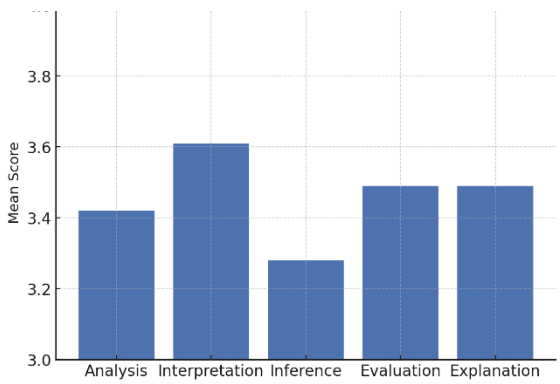

Figure 2: Average Scores in Critical Thinking Dimensions

4.4 Correlation Analysis

Pearson's correlation coefficients revealed a significant negative association between AI dependence and overall critical thinking ($r = -.36$, $p < .01$). At the dimensional level, inference ($r = -.41$, $p < .01$) and evaluation ($r = -.33$, $p < .01$) were the most strongly correlated with AI dependence. Table 5 presents the correlation matrix, and Figure 3 illustrates the negative linear relationship between AI dependence and total critical thinking.

Table 5: Correlation AI dependence and critical thinking dimensions

Variable	r de Pearson	IC 95%	p	ρ de Spearman	p	r parcial[1]	p
Critical Thinking	-.36	[−0.48, −0.23]	< .01	−0.35	< .01	−0.33	< .01
Analysis	-.29	[−0.42, −0.15]	< .01	−0.28	< .01	−0.27	< .01
Interpretation	-.21	[−0.34, −0.07]	.004	−0.19	.006	−0.20	.008
Inference	-.41	[−0.52, −0.29]	< .01	−0.40	< .01	−0.36	< .01
Evaluation	-.33	[−0.45, −0.20]	< .01	−0.31	< .01	−0.31	< .01
Explanation	-.27	[−0.40, −0.13]	< .01	−0.26	< .01	−0.24	.001

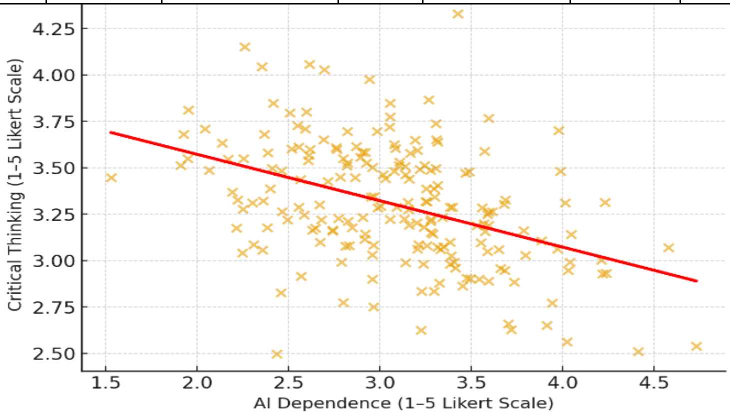

Figure 3: Scatterplot of AI Dependence and Critical Thinking

4.5 Regression Analysis

Hierarchical models included gender and prior GPA (z-scored) as controls. Model 1 (controls): $R^2 = .06$, Adjusted $R^2 = .05$, $F(2,197) = 6.30$, $p = .002$ (GPA positive; gender ns). Model 2 (+ AI dependence): $R^2 = .21$, Adjusted $R^2 = .20$, $F(3,196) = 17.47$, $p < .001$; $\Delta R^2 = .15$, $p < .001$. AI dependence negatively predicted CT total ($\beta = -.39$, 95% CI [−.51, −.27], $p < .001$), with semi-partial $r^2 = .13$ (medium practical effect). Assumptions held (linearity, normal residuals, homoscedasticity), and multicollinearity was low (VIF < 3) (see Table 6).

At the dimension level (same controls), effects were strongest for Inference ($\beta = -.34$, 95% CI [−.46, −.23], $p < .001$; Adjusted $R^2 = .22$) and Evaluation ($\beta = -.26$, 95% CI [−.38, −.14], $p < .001$; Adjusted $R^2 = .16$), followed by Analysis ($\beta = -.24$, $p < .001$), Explanation ($\beta = -.21$, $p = .002$), and Interpretation ($\beta = -.17$, $p = .018$) (see Table 7).

Table 6: Hierarchical regression predicting critical thinking

Predictor	Model 1 (Controls) β [95% CI]	p	Model 2 (+ AI Dependence) β [95% CI]	p
Gender (female = 1)	0.05 [−0.06, 0.16]	.372	0.04 [−0.06, 0.15]	.437
Prior GPA (z)	0.22 [0.08, 0.36]	.002	0.17 [0.05, 0.30]	.006
AI dependence (z)	—	–	−0.39 [−0.51, 0.27]	<.001
Model fit	R^2 = .06; Adj. R^2 = .05		R^2 = .21; Adj. R^2 = .20; ΔR^2 = .15	—
F, df	F(2, 197) = 6.30	.002	F(3, 196) = 17.47	<.001
Semi-partial r^2 (AI Dep.)	—	–	.13	—
N	200		200	

Notes. CT Total and predictors standardized except gender (0 = male, 1 = female). Robustness checks: model assumptions met; VIF < 3.

Table 7: Dimension-level regressions predicting CT dimensions

Outcome (CT dimension)	β (AI Dependence) [95% CI]	p	Adjusted R^2	N
Inference	−0.34 [−0.46, −0.23]	<.001	.22	200
Evaluation	−0.26 [−0.38, −0.14]	<.001	.16	200
Analysis	−0.24 [−0.36, −0.12]	<.001	.12	200
Explanation	−0.21 [−0.34, −0.08]	.002	.10	200
Interpretation	−0.17 [−0.31, −0.03]	.018	.07	200

4.6 Summary of Findings

Overall, AI dependence was moderate, and critical thinking exhibited an uneven profile, with Inference weakest. Correlation and regression results indicated medium-sized negative associations between AI dependence and CT—particularly for Inference and Evaluation—that hold after controlling for gender and prior GPA (see Tables 6 and 7), consistent with the scaffolding vs. substitution account and the Peruvian context

5. Discussion

First-year students in Trujillo exhibited moderate AI dependence alongside an uneven critical-thinking (CT) profile, with Inference weakest and Interpretation/Explanation comparatively stronger. Correlation and hierarchical regression yielded medium-sized negative associations between AI dependence and CT that remained after controlling for gender and prior GPA, with Adjusted R^2 reaching .22 and a semi-partial $r^2 \approx .13$ for the focal predictor. Effects concentrate on Inference and Evaluation, but not uniformly across dimensions, indicating a selective link to higher-order reasoning (see Tables 6–7). This pattern resonates with international warnings about overreliance while grounding them in a Peruvian cohort (Kasneci et al., 2023; Rasul et al., 2023). Overall, the results portray widespread but largely support-oriented AI use, and they suggest that dependence is most strongly associated with the stages where evidence is weighed, warrants are articulated, and conclusions must be justified.

Theoretical lenses clarify these associations. In Facione's Delphi model, Inference and Evaluation anchor reflective judgment, aligning with their stronger links to the AI dependence observed here (Facione, 2015). From Zimmerman's self-regulated learning perspective, students appear to offload monitoring and evaluation, preserving task comprehension while outsourcing effortful steps (Malik et al, 2024; Zimmerman, 2002).

In Salomon's terms, frequent, goal-substituting use turns AI into a substitute rather than a scaffold (Jose et al.2025). In Vygotsky's view, without guided internalization, support plateaus as dependence, especially among novices (Liu et al., 2024). The deep vs. surface distinction helps explain why substitution maps onto shallower engagement and weaker warranting of claims (Saharuddin et al., 2024). Together, these lenses predict precisely the selective pattern we observed.

Against international research, the findings both converge and nuance previous evidence. Studies warning that overreliance threatens autonomy are mirrored in our metacognitive displacement pattern, in which monitoring and evaluation are the most affected stages (Grájeda et al., 2024; Umbase, 2023; Rasul et al., 2023). Similar risks reported in the UAE suggest structural features of AI use, not only local artifacts (Medina et al., 2024).

The low Inference scores accord with concerns about superficial reasoning and premature closure (Luciano, 2024), while comparatively higher Interpretation/Explanation aligns with reports that AI can aid comprehension and communication when used for support (Lo, 2023). By isolating dimension-specific associations and quantifying their magnitude, our study adds precision to global debates about how, where, and for whom dependence matters most.

Within Latin America, this study complements and extends prior work by foregrounding context. In Peru, faculty adoption of ChatGPT is shaped by pedagogical knowledge, and structured integration can improve students' research skills (Ríos Gonzales et al., 2024; 2025). Our results illuminate the other side: unstructured, substitution-oriented use among novices is negatively associated with higher-order CT, especially Inference/Evaluation, underscoring the role of pedagogical mediation and institutional guidance (Nabavi & Farajollahi, 2024).

Regional conditions—digital divides, uneven training, and incipient policies—likely increase substitution patterns, explaining the stronger links we document. Thus, the contribution is twofold: localized evidence from an underrepresented setting, and a framework for interpreting how contextual constraints modulate the balance between scaffolding and substitution in early university cohorts.

Implications follow for design and policy. To shift from substitution to scaffolding, instructors can require prompt deconstruction and inference audits targeting Inference, source triangulation and justification rubrics targeting Evaluation, and AI-as-peer debates to reconstruct reasoning. Institutions should institutionalize critical digital literacies, positioning AI as a partner to be verified rather than a replacement. Limitations include a non-probability sample at a private university, cross-sectional data that preclude causal inference, and self-report measures.

Future research should test longitudinal change and experimental interventions to evaluate whether scaffolded use attenuates the negative associations and conduct multi-site Latin American studies to estimate contextual moderators (Kasneci et al., 2023; Ramírez-Montoya, 2022; Rasul et al., 2023). Together, these steps can align adoption with higher-order learning goals.

6. Conclusion

First-year students in Trujillo display moderate AI dependence and an uneven critical-thinking profile, with Inference and Evaluation most vulnerable while Interpretation and Explanation remain comparatively intact. This selective pattern aligns with frameworks by Facione, Zimmerman, Salomon, and Vygotsky, suggesting that unmediated AI use can shift from scaffold to substitute, hindering the internalization of reasoning strategies.

For policy-makers, we recommend codifying institutional standards for AI-supported work: (a) minimum verification and source-triangulation requirements; (b) audit trails (prompt/output logs attached to submissions); (c) assessments that grade the reasoning process (not just products); (d) academic-integrity protocols that explicitly cover generative tools; and (e) investments in faculty development, critical digital-literacy curricula, and equitable access to vetted AI tools and reference databases.

Methodological implications. Our non-probability sample from a private university, cross-sectional design, and self-report measures limit external validity and raise common-method concerns. Future research should use probability, multi-site sampling (public/private, multi-city), longitudinal panels, and pre-registered experiments that vary AI-scaffolding requirements. To reduce self-report bias, they should include behavioral/trace data (prompt logs, version histories), performance-based CT assessments, and mixed methods (e.g., think-alouds). Analytically, they should test measurement invariance, employ multi-level models for course/instructor effects, and examine mediators/moderators (purpose of use, training exposure, integrity norms).

7. Conflict of Interest

The authors declare that there are no financial, commercial, institutional, or personal conflicts of interest that could have influenced the conception, conduct, analysis, or reporting of this study. Specifically, the authors have no economic ties or contractual commitments with providers of artificial intelligence tools mentioned in the manuscript, and the interpretation of the results is independent of any external interests. The study did not receive funding that conditioned its methodological decisions or its conclusions.

8. Acknowledgments

We thank the students who participated in the study and the academic and administrative staff who facilitated data collection. We are also grateful to the anonymous peer reviewers for their comments, which helped improve the clarity and rigor of the manuscript. The authors acknowledge that ChatGPT was used as the sole artificial intelligence tool from the outset of the manuscript, exclusively to assist with language and grammatical refinement. The scientific content, methodological design, data analysis, and conclusions are the sole responsibility of the authors.

9. References

Adiguzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology*, *15*(3), Article 429. https://doi.org/10.30935/cedtech/13152

Aldawsari, H. K., Shahat, H. A., Al-Ruwaili, H. A., Gaber, S. A., Aldughaysh, A. M., & Mohmed, T. A. (2023). Developing productive thinking skills in the field of artistic works using the SCAMPER strategy for twice exceptional students. *International Journal of Learning, Teaching and Educational Research, 22*(12), 1–20. https://doi.org/10.26803/ijlter.22.12.1

Almulla, M., & Ali, S. I. (2024). The changing educational landscape for sustainable online experiences: Implications of ChatGPT in Arab students' learning experience. *International Journal of Learning, Teaching and Educational Research*, *23*(9), 285–306. https://doi.org/10.26803/ijlter.23.9.15

Alnasib, B. N. M. (2023). Factors affecting faculty members' readiness to integrate artificial intelligence into their teaching practices: A study from the Saudi higher education context. *International Journal of Learning, Teaching and Educational Research*, *22*(8), 465–491. https://doi.org/10.26803/ijlter.22.8.24

Alrishan, A. M. (2023). Determinants of intention to use ChatGPT for professional development among Omani EFL pre-service teachers. *International Journal of Learning, Teaching and Educational Research*, *22*(12), 187–209. https://doi.org/10.26803/ijlter.22.12.10

Ampo, W. M.G., Ayuban, A. R., Avellaneda, S.L. A., & Go, D.T. (2025). Exploring teachers' lived experiences in integrating ChatGPTin classroom practices. *International Journal of Education and Emerging Practices, 1*(1), 17-28. https://doi.org/10.63236/injeep.1.1.2

Anand, M. D. (2024). *Research Methodology: Methods and Techniques*. SAN Publications. https://doi.org/10.59646/rmmt/171

Atchley, P., Pannell, H., Wofford, K., Hopkins, M., & Atchley, R. A. (2024). Human and AI collaboration in the higher education environment: Opportunities and concerns. *Cognitive Research: Principles and Implications*, *9*(20). https://doi.org/10.1186/s41235-024-00547-9

Banihashem, S. K., Taghizadeh Kerman, N., Noroozi, O., Moon, J., & Drachsler, H. (2024). Feedback sources in essay writing: Peer-generated or AI-generated feedback? *International Journal of Educational Technology in Higher Education*, *21*(23). https://doi.org/10.1186/s41239-024-00455-4

Bitzenbauer, P. (2023). ChatGPT in physics education: A pilot study on easy-to-implement activities. *Contemporary Educational Technology*, *15*(3), ep430. https://doi.org/10.30935/cedtech/13176

Bukar, U. A., Sayeed, M. S., Abdul Razak, S. F., Yogarayan, S., & Sneesl, R. (2024). Prioritizing ethical conundrums in the utilization of ChatGPT in education through an analytical hierarchical approach. *Education Sciences*, *14*(9), 959. https://doi.org/10.3390/educsci14090959

Castagnola Rossini, G., Urbano Gutiérrez, L. P., & Pérez Azahuanche, M. A. (2025). La inteligencia artificial para desarrollar las habilidades investigativas en docentes universitarios [Artificial intelligence to develop research skills in university professors]. *Horizontes. Revista de Investigación en Ciencias de la Educación*, *9*(37), 1009–1026. https://doi.org/10.33996/revistahorizontes.v9i37.965

Celik, I. (2023). Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. *Computers in Human Behavior*, *138*, 107468. https://doi.org/10.1016/j.chb.2022.107468

Chamorro-Atalaya, O., Huarcaya-Godoy, M., Durán-Herrera, V., Nieves-Barreto, C., Suarez-Bazalar, R., Cruz-Telada, Y., Alarcón-Anco, R., Huayhua-Mamani, H., Vargas-Diaz, A., & Balarezo-Mares, D. (2023). Application of the chatbot in university education: A systematic review on the acceptance and impact on learning. *International Journal of Learning, Teaching and Educational Research*, *22*(9), 156–178. https://doi.org/10.26803/ijlter.22.9.9

Cohen, L., Manion, L., & Morrison, K. (2018). *Research Methods in Education* (8th ed.). Routledge. https://doi.org/10.4324/9781315456539

Creswell, J. W., & Creswell, J. D. (2023). Research design: Qualitative, quantitative, and mixed methods approach (6th ed.). SAGE. https://www.ucg.ac.me/skladiste/blog_609332/objava_105202/fajlovi/Creswell.pdf

De Jesus, F. S., Ibarra, L. M., Pasion, B. J., Villanueva, W., & Leyesa, M. (2024). ChatGPT as an artificial intelligence learning tool for Business Administration students in Nueva Ecija, Philippines. *International Journal of Learning, Teaching and Educational Research*, *23*(6), 348–372. https://doi.org/10.26803/ijlter.23.6.16

Dergaa, I., Chamari, K., Zmijewski, P., & Saad, H. B. (2023). From human writing to artificial intelligence generated text: Examining the prospects and potential threats of ChatGPT in academic writing. *Biology of Sport*, *40*(2), 615–622. https://doi.org/10.5114/biolsport.2023.125623

Duong, C. D. (2024). ChatGPT adoption and digital entrepreneurial intentions: Empirical research based on the theory of planned behaviour. *Entrepreneurial Business and Economics Review*, *12*(2), 129–142. https://doi.org/10.15678/EBER.2024.120208

Essien, A., Bukoye, O. T., O'Dea, X., & Kremantzis, M. (2024). The influence of AI text generators on critical thinking skills in UK business schools. *Studies in Higher Education*, *49*(5), 865–882. https://doi.org/10.1080/03075079.2024.2316881

Facione, P. A. (2015). *Critical thinking: What it is and why it counts* (Rev. ed.). Insight Assessment. https://www.insightassessment.com

Flores-Vivar, J. M., & García-Peñalvo, F. J. (2023). Reflexiones sobre la ética, potencialidades y retos de la Inteligencia Artificial en el marco de la Educación de Calidad (ODS4). *Comunicar, 31*(74), 37–47. https://doi.org/10.3916/C74-2023-03

Fuchs, K. (2023). Exploring the opportunities and challenges of NLP models in higher education: Is ChatGPT a blessing or a curse? *Frontiers in Education*, *8*, Article 1166682. https://doi.org/10.3389/feduc.2023.1166682

Funda, V., & Mbangeleli, N. B. A. (2024). Artificial Intelligence (AI) as a tool to address academic challe nges in South African higher education. *International Journal of Learning, Teaching and Educational Research*, *23*(11), 1–15. http://dx.doi.org/10.26803/ijlter.23.11.27

Goh, P. S., & Sandars, J. (2024). Application of ChatGPT-assisted problem-based learning teaching method in clinical medical education. *BMC Medical Education*, *24*(1), 1-10. https://doi.org/10.1186/s12909-024-04812-5

Gouia-Zarrad, R., & Gunn, C. (2024). Enhancing students' learning experience in mathematics class through ChatGPT. *International Electronic Journal of Mathematics Education*, *19*(3), em0781. https://doi.org/10.29333/iejme/14614

Grájeda, A., Córdova, P., Córdova, J. P., Laguna-Tapia, A., Burgos, J., Rodríguez, L., Arandia, M., & Sanjinés, A. (2024). Embracing artificial intelligence in the arts classroom: Understanding student perceptions and emotional reactions to AI tools. *Cogent Education*, *11*(1), 2378271. https://doi.org/10.1080/2331186X.2024.2378271

Guo, Y., & Lee, D. (2023). Leveraging ChatGPT for enhancing critical thinking skills. *Journal of Chemical Education*, *100*(11), 4876–4883. https://doi.org/10.1021/acs.jchemed.3c00505

Han, X., Wang, J., & Liu, Q. (2024). A qualitative analysis of Chinese higher education students' intentions and influencing factors in using ChatGPT: A grounded theory approach. *Scientific Reports*, *14*(1), 1-12. https://doi.org/10.1038/s41598-024-50748-3

Iqbal, Z., & Rahman, S. (2024). Adoption and impact of ChatGPT in computer science education: A case study on a database administration course. *AI Switzerland*, *6*(3), 215–231. https://doi.org/10.3390/aiswitzerland6030015

Jose, B., Cherian, J., Verghis, A.M., Varghise, S.M., & Joseph, S. (2025). The cognitive paradox of AI in education: between enhancement and erosion. *Frontiers in Psychology*,*16,* 1550621. https://doi.org/10.3389/fpsyg.2025.1550621

Kasneci, E., Sessler, K., Kitchenham, A., Kasneci, G., & Bannert, M. (2023). ChatGPT and other large language models: Applications in education, opportunities, and challenges. *Computers and Education: Artificial Intelligence*, *4*, 100154. https://doi.org/10.1016/j.caeai.2023.100154

Kooli, C. (2023). ChatGPT and the future of education: A critical analysis of the potential benefits and risks. *Journal of Applied Learning & Teaching*, *6*(1), 1–12. https://doi.org/10.37074/jalt.2023.6.1.9

Liu, X., Chen, J., & Wang, Y. (2024). Artificial intelligence and the transformation of higher education institutions: A systems approach. *Sustainability Switzerland*, *16*(5), 3456–3471. https://doi.org/10.3390/su16053456

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education and Information Technologies*, *28*(7), 9011–9034. https://doi.org/10.1007/s10639-023-11870-8

Luciano, R. G. (2024). The dual impact of ChatGPT on learning and ethics among Bachelor of Science in Information Technology (BSIT) students. *International Journal of Learning, Teaching and Educational Research*, *23*(12), 78–95. https://doi.org/10.26803/ijlter.23.12.5

Malik, M. A., Amjad, A. I., & Aslam, S. (2024). Global insights: ChatGPT's influence on academic and research writing, creativity, and plagiarism policies. *Frontiers in Research Metrics and Analytics*, *9*, Article 1486832. https://doi.org/10.3389/frma.2024.1486832

Medina, W. A., Flores Velásquez, C. H., Olivares Zegarra, S. del R., Morales Romero, G. P., Quispe Andía, A., Aybar Bellido, I. E., Palacios Huaraca, C. R., Aldana Trejo, F. H., & Bravo, N. A. (2024). Using ChatGPT in university academic writing: A bibliometric review study on the implications for writing reports, papers, essays, and theses. *International Journal of Learning, Teaching and Educational Research*, *23*(12), 360–381. https://doi.org/10.26803/ijlter.23.12.19

Nabavi, R., & Farajollahi, M. (2024). Challenges and opportunities of generative AI for higher education as explained by ChatGPT. *Education Sciences*, *14*(2), 287–299. https://doi.org/10.3390/educsci14020287

OECD. (2024). *Education at a Glance 2020: OECD Indicators*. OECD Publishing. https://doi.org/10.1787/69096873-en

Ramírez-Montoya, M. S., Castillo-Martínez, I. M., Sanabria-Z, J., & Miranda, J. (2022). Complex thinking in the framework of Education 4.0 and open innovation—A systematic literature review. *Journal of Open Innovation: Technology, Market, and Complexity, 8*(1), 4. https://doi.org/10.3390/joitmc8010004

Rasul, T., Nair, S., Kalendra, D., Robin, M., de Oliveira Santini, F., Ladeira, W. J., Sun, M., Day, I., & Rather, R. A. (2023). The role of ChatGPT in higher education: Benefits, challenges, and future research directions. *Journal of Applied Learning & Teaching*, *6*(1). https://doi.org/10.37074/jalt.2023.6.1.29

Ríos Gonzales, J. D., Tomanguilla Reyna, J., Vereau Amaya, E. A., & Vásquez Luján, I. G. (2025). Evaluation of the impact of ChatGPT on the development of research skills in higher education. *International Journal of Learning, Teaching and Educational Research*, *24*(4), 370–390. https://doi.org/10.26803/ijlter.24.4.18

Saharuddin, M. H., Nasir, M. K. M., & Mahmud, M. S. (2024). Exploring teachers' technological pedagogical content knowledge in utilising artificial intelligence (AI) for teaching. *International Journal of Learning, Teaching and Educational Research*, *24*(1), 108–123. https://doi.org/10.26803/ijlter.24.1.7

Umbase, R. S. (2023). Implementing Technological Pedagogical and Content Knowledge from the Social Studies Learning Management Perspective. *International Journal of Learning, Teaching and Educational Research*, *22*(11), 401–418. https://doi.org/10.26803/ijlter.22.11.21

Yue, M., Jong, M. S.-Y., & Ng, D. T. K. (2024). Understanding K–12 teachers' technological pedagogical content knowledge readiness and attitudes toward artificial intelligence education. *Education and Information Technologies*. *29*, 19505–19536 https://doi.org/10.1007/s10639-024-12621-2

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, *41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2