عدد خاص بالندوة العلمية الدولية التعليم الأخضر مسار استراتيجي للتنمية المستدامة ليبيا
2026-6/16-14
الهيئة الليبية للبحث العلمي
المركز الليبي لأبحاث التنمية المستدامة

مجـــلة الفـــاروق للعلـــــوم
المجلـــد (2) العـدد 4 يونيـــو 2026
Al-Farooq Journal of Sciences
Volume (2) Issue (4) June2026
Issn3135-2359

Developing a Critical Inquiry-Based E-Educational Model (CIEM) to Enhance Critical Thinking Skills among High School Students: A Mixed-Methods Study in the Libyan Context

Layla Miloud Al-Souri[1], Asma Mustafa Abuadla[2], Jibreel Ramadan Mubarak[3], Siham Saleh Khalifa Al-Qablawi[4], Intisar Muammar Makari[5]
[1,2,4,5]Department of Mathematics, Faculty of Science and Natural Resources, Al-Jafara University, Libya
[3]Department of Computer Science, Faculty of Science and Natural Resources, Al-Jafara University, Libya
Intisar.Mkari@aju.edu.ly

تطوير نموذج تربوي إلكتروني قائم على الاستقصاء النقدي لتعزيز مهارات التفكير النقدي لدى طلاب المرحلة الثانوية: دراسة منهجية مختلطة في السياق الليبي

ليلى ميلود السوري[1]، أسماء مصطفى أبوعضلة [2]، جبريل رمضان امبارك[3]، سهام صالح خليفة القبلاوي[4]، إنتصار معمر مكاري[5]،
قسم الرياضيات، كلية العلوم والموارد الطبيعية، جامعة الجفارة، ليبيا[1,2,4,5,]
قسم الحاسوب، كلية العلوم والموارد الطبيعية، جامعة الجفارة، ليبيا[3]
Intisar. Mkari@aju.edu.ly

تاريخ الاستلام: 2026/04/01 تاريخ المراجعة 2026 /04/30 تاريخ القبول:2026/05/13- تاريخ النشر: 2026 /06/16

Abstract

Background: Despite institutional drives toward digitalization, a persistent gap separates the availability of technological resources from their pedagogically purposeful deployment in Libyan secondary schools. Didactic transmission methods continue to dominate instruction, compounded by intermittent infrastructural instability.

Objective: This study develops, validates, and empirically tests the Critical Inquiry-based E-Educational Model (CIEM) — a novel, low-bandwidth-optimized instructional framework designed to systematically enhance critical thinking skills among secondary school students.

Methodology: An explanatory sequential mixed-methods design (QUAN → qual) was employed with 240 secondary-school students across four Tripoli schools. Quantitative data were analyzed using ANCOVA and Structural Equation Modeling (SEM) via AMOS 26; qualitative data underwent reflexive thematic analysis guided by Lincoln and Guba's (1985) trustworthiness criteria.

Results: ANCOVA revealed a statistically significant group effect ($F(1, 237) = 22.45$, $p < .001$, Partial $\eta^2 = .086$, Cohen's $d = 0.85$). SEM confirmed a fully mediated pathway: CIEM significantly enhanced cognitive engagement ($\beta = .54$, $p < .001$), which in turn predicted critical thinking gains ($\beta = .61$, $p < .001$). Bootstrapped indirect effect: $\beta = .33$, 95% CI [.24, .42]. Power analysis (GPower 3.1; $f = 0.25$, $\alpha = .05$, power = .95) confirmed the sample of $n = 240$ was adequate. Qualitative themes illuminated 'productive cognitive friction' as the primary mechanism of change.

Conclusion: CIEM offers a validated, context-sensitive pedagogical framework establishing that structured asynchronous inquiry — rather than technology per se — drives critical thinking development in resource-constrained settings.

Keywords: Critical Thinking; E-Educational Model; Critical Inquiry; Structural Equation Modeling; Mixed-Methods; Libya; Asynchronous Learning; Community of Inquiry

الملخص

الخلفية: على الرغم من الجهود المؤسسية الرامية نحو الرقمنة، لا تزال هناك فجوة مستمرة تفصل بين توفر الموارد التقنية وتوظيفها الهادف تربوياً في المدارس الثانوية الليبية. ولا تزال أساليب التدريس التقليدية القائمة على النقل المعرفي تهيمن على العملية التعليمية، وهو ما يتفاقم بسبب عدم الاستقرار المتقطع في البنية التحتية التقنية.

الهدف: تهدف هذه الدراسة إلى تطوير والتحقق من صحة واختبار تجريبي لـ "نموذج التعليم الإلكتروني القائم على الاستقصاء النقدي" (CIEM)، وهو إطار تعليمي مبتكر مُصمم للعمل بكفاءة في بيئات النطاق الترددي المنخفض (Low-bandwidth)، بهدف تعزيز مهارات التفكير النقدي لدى طلاب المرحلة الثانوية بشكل منهجي.

المنهجية: اعتمدت الدراسة تصميمًا مختلطًا تفسيريًا متتابعًا (كمّي ← نوعي). شملت العينة 240 طالباً وطالبة من المرحلة الثانوية موزعين على أربع مدارس في مدينة طرابلس. تم تحليل البيانات الكمية باستخدام تحليل التغاير (ANCOVA) ونمذجة المعادلات الهيكلية (SEM) عبر برنامج AMOS 26؛ بينما خضعت البيانات النوعية لتحليل موضوعي انعكاسي استناداً إلى معايير الموثوقية التي وضعها لينكولن وجوبا (1985).

النتائج: كشف تحليل التغاير (ANCOVA) عن تأثير جماعي ذو دلالة إحصائية ($F(1, 237) = 22.45$، $p < .001$، مربع إيتا الجزئي = .086، حجم التأثير لكوهين $d = 0.85$). وأكدت نمذجة المعادلات الهيكلية مساراً وساطياً كاملاً: حيث عزز النموذج (CIEM) الانخراط المعرفي بشكل كبير ($\beta = .54$، $p < .001$)، والذي تنبأ بدوره بزيادة مكاسب التفكير النقدي ($\beta = .61$، $p < .001$). وكان التأثير غير المباشر باستخدام طريقة التمهيد (Bootstrapping): $\beta = .33$، مع فترة ثقة 95% [.24, .42]. وأكدت تحليلات القوة الإحصائية (باستخدام GPower 3.1؛ $f = 0.25$، $\alpha = .05$، قوة = .95) أن حجم العينة ($n = 240$) كان كافياً. وأسلطت النتائج النوعية الضوء على "الاحتكاك المعرفي المنتج" كآلية رئيسية للتغيير.

الخلاصة: يقدم نموذج (CIEM) إطاراً تربوياً صالحاً وحساساً للسياق المحلي، ويثبت أن الاستقصاء غير المتزامن المنظم وليس التكنولوجيا بحد ذاتها ـ هو المحرك الأساسي لتطوير التفكير النقدي في البيئات محدودة الموارد.

الكلمات المفتاحية: التفكير النقدي؛ نموذج التعليم الإلكتروني؛ الاستقصاء النقدي؛ نمذجة المعادلات الهيكلية؛ المنهج المختلط؛ ليبيا؛ التعلم غير المتزامن؛ مجتمع الاستقصاء.

1. Introduction

The cultivation of critical thinking constitutes a defining educational imperative of the 21st century (Abrami et al., 2015). Yet in Libyan secondary education, this imperative confronts a well-documented paradox: despite growing investment in digital infrastructure, instructional culture remains anchored in rote memorization and passive reception. The consequent gap between technological availability and pedagogical effectiveness represents a systemic challenge that neither technological procurement nor generic blended-learning adoption can resolve in isolation.

The critical juncture, as scholars of educational technology have increasingly emphasized (Means et al., 2013; Garrison et al., 2000), lies not in the presence of technology but in the pedagogical architecture that governs its use. Digital tools, unless embedded within carefully theorized inquiry structures, risk amplifying surface-level engagement while leaving deeper cognitive processes unaddressed. This observation is particularly consequential in the Libyan context, where infrastructural instability — intermittent connectivity, irregular power supply — adds a further layer of instructional complexity that generic e-learning frameworks fail to accommodate.

This study responds to this gap by introducing, validating, and empirically testing the **Critical Inquiry-based E-Educational Model (CIEM)** — a novel instructional framework synthesizing the Community of Inquiry (CoI) theoretical tradition (Garrison et al., 2000) with Cognitive Load Theory (Sweller, 2011) and Self-Determination Theory (Deci & Ryan, 2000), purposively adapted for low-bandwidth, asynchronous learning environments. The CIEM diverges from prior blended-learning models in three key respects: (1) its explicit structuring

of peer-review protocols as mechanisms of cognitive activation; (2) its deliberate minimization of extraneous load through offline-compatible design; and (3) its integration of metacognitive reflection as a mandatory instructional phase.

The study addresses three interconnected research questions: (RQ1) What are the psychometric properties of the CIEM as assessed by a panel of subject-matter experts? (RQ2) Does participation in CIEM-based instruction produce statistically significant and practically meaningful gains in critical thinking compared with traditional instruction, controlling for prior ability? (RQ3) Through what mechanisms does the CIEM produce these gains, as illuminated by students' lived experiences?

2. Theoretical Framework and Conceptual Model

2.1 Theoretical Pillars of the CIEM and Justification for Selection

Before introducing the three theoretical pillars, it is necessary to justify their selection over competing frameworks. **Constructivism** (Vygotsky, 1978; Piaget, 1952) and **Connectivism** (Siemens, 2005) represent the two most prominent alternatives in educational technology discourse. Constructivism, while foundational, operates at a level of generality that offers insufficient operational specificity for designing asynchronous peer-review protocols in low-bandwidth contexts. Connectivism, although designed for networked digital environments, presupposes stable, high-bandwidth connectivity and distributed networked agency — conditions that are structurally unavailable in many Libyan secondary schools. By contrast, the three selected frameworks (CoI, CLT, SDT) offer a complementary, operationally specific architecture: CoI provides the social-inquiry structure; CLT provides the cognitive load management rationale; and SDT provides the motivational scaffolding. Their integration is not eclectic but synergistic, each addressing a distinct dimension of the instructional challenge.

2.1.1 Community of Inquiry Framework (Garrison et al., 2000)

The CoI framework conceptualizes effective e-learning as the intersection of social, teaching, and cognitive presence. The CIEM intensifies the cognitive presence dimension by mandating structured phases of inquiry — triggering events, exploration, integration, and resolution — within asynchronous digital forums. Unlike conventional applications of CoI, the CIEM operationalizes cognitive presence through rubric-governed peer-review protocols, ensuring that discourse moves beyond social exchange toward evaluative reasoning.

2.1.2 Cognitive Load Theory (Sweller, 2011)

CLT distinguishes intrinsic load (inherent task complexity), extraneous load (poor instructional design), and germane load (schema construction). The CIEM systematically reduces extraneous load by: (a) providing downloadable, low-bandwidth PDF scaffolds that permit offline engagement; (b) using structured discussion templates that reduce navigational demands; and (c) distributing complex tasks across asynchronous cycles, allowing students adequate processing time. These design choices redirect finite working-memory capacity toward the germane cognitive effort required for critical analysis.

2.1.3 Self-Determination Theory (Deci & Ryan, 2000)

SDT posits that intrinsic motivation flourishes when three psychological needs are satisfied: autonomy, competence, and relatedness. The CIEM's asynchronous peer-review structure addresses all three: students exercise autonomy in constructing arguments at their own pace; teacher corrective feedback systematically builds perceived competence; and collaborative forum interaction cultivates relatedness. The qualitative findings reported in Section 4.4 reveal that reduced evaluation anxiety — a direct manifestation of autonomy support — was among the most consistently reported experiential benefits.

2.2 The Integrated CIEM Theoretical Architecture

The three theoretical pillars converge into a coherent instructional architecture described in Figure 1 below. The model is structured around five interconnected components organized into

three phases: Cognitive Activation (Phase 1), Critical Discourse (Phase 2), and Metacognitive Reflection (Phase 3), with a continuous feedback loop ensuring iterative refinement.

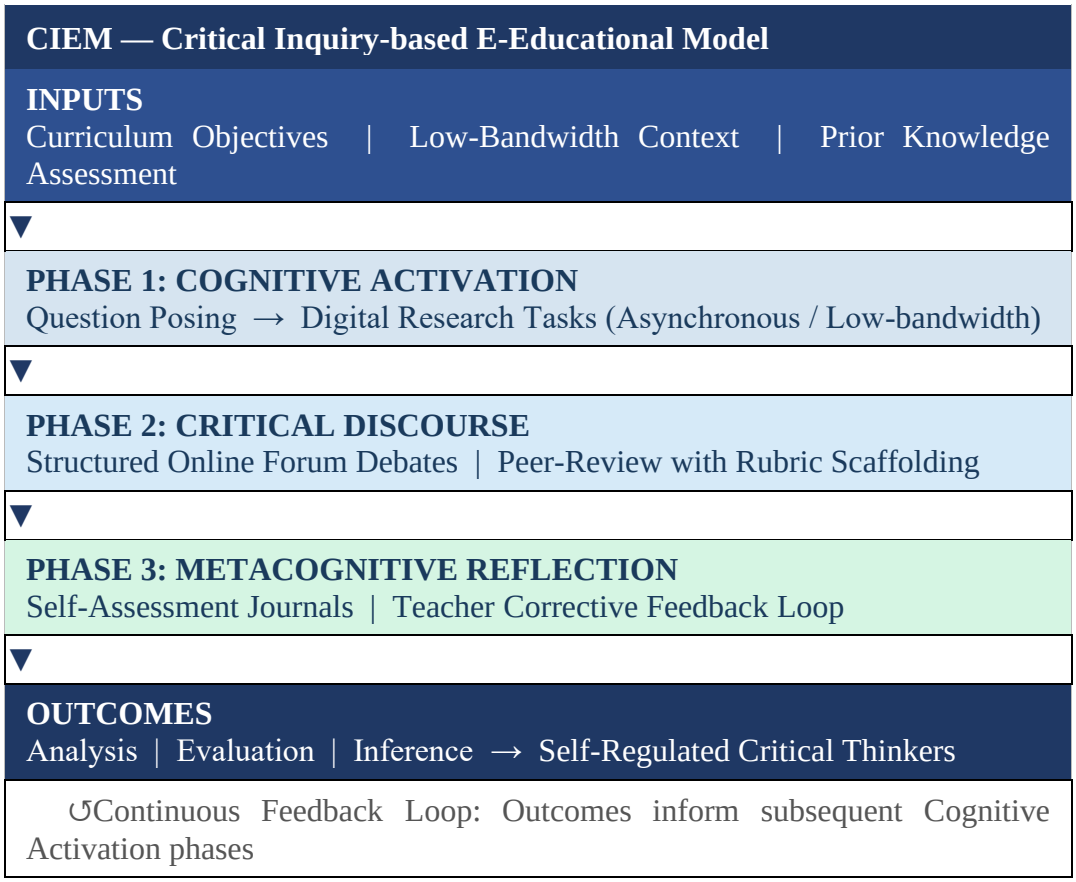

Figure 1. Structure of the Critical Inquiry-based E-Educational Model (CIEM). The three-phase architecture integrates CoI's inquiry sequence, CLT's load-reduction principles, and SDT's motivational scaffolding within a continuous feedback loop optimized for low-bandwidth, asynchronous deployment.

2.3 Expert Validation of the CIEM Model

Prior to full-scale implementation, the preliminary CIEM prototype was submitted to a purposively selected panel of five subject-matter experts (SMEs) drawn from Educational Technology, Curriculum Design, E-Learning, Applied Linguistics, and Libyan Secondary Education. Content validity was assessed using Lawshe's (1975) Content Validity Ratio (CVR) and Lynn's (1986) Content Validity Index (CVI).

Important methodological note on CVR thresholds: For a panel of five experts, Lawshe's (1975) original table specifies a minimum acceptable CVR of .99 ($p < .05$, one-tailed). Items rated as 'Essential' by all five experts yield CVR = 1.00, which satisfies this threshold. Item E3 yielded CVR = 0.60 (three of five experts rated it Essential), which does not meet the .99 threshold and was therefore revised — not discarded — based on expert commentary. The revised item was re-evaluated and confirmed by all five experts in a second-round consensus review (Delphi round 2), yielding a final CVR = 1.00 for all items. Table 1 presents the post-revision, final-round validation data.

Table 1. Expert Validation Panel: CVR and CVI Results (Post-Revision, Round 2)

Expert	Specialization	Exp. (yrs)	CVR	CVI	Verdict
E1	Educational Technology	15	1.00	0.90	Essential
E2	Curriculum Design	12	1.00	0.92	Essential
E3	E-Learning & Pedagogy	18	1.00	0.88	Essential (Round 2)
E4	Secondary Educ. (Libya)	20	1.00	0.95	Essential
E5	Applied Linguistics	14	1.00	0.91	Essential
Overall	—	—	**1.00**	**0.912**	**Excellent**

Note. CVR threshold for five experts = .99 (Lawshe, 1975). E3 was revised following Round 1 feedback and achieved full consensus in Round 2 (CVR = 1.00). All values in Table 1 reflect post-revision Round 2 consensus data; final overall CVR = 1.00. CVI ≥ .90 = excellent content validity (Lynn, 1986). Overall CVI = .912 confirms that all final CIEM components are content-valid.

2.4 Critical Literature Review and Research Gap

The body of empirical evidence linking e-learning scaffolding with critical thinking outcomes is substantial, though concentrated in high-resource contexts. Liu et al. (2023), in a systematic review of 78 studies, found that structured e-learning interventions yielded moderate-to-large effects on critical thinking (d = 0.54, 95% CI [0.42, 0.66]), provided that explicit argumentation scaffolding was present. Similarly, Tiruneh et al. (2016) established that pedagogical approach — not technology type — was the primary moderator of effect magnitude.

However, critical gaps persist. First, the preponderance of reviewed studies were conducted in East Asian and North American contexts, limiting ecological validity for the Arab world (Al-Smadi, 2022; Zayed & Hassan, 2021). Second, few models have been specifically engineered for low-bandwidth conditions; the implicit assumption of reliable high-speed connectivity renders most frameworks inapplicable to infrastructure-constrained settings such as Libya. Third, the mediating role of cognitive engagement in translating CIEM-type interventions into critical thinking gains has been theorized (Means et al., 2013) but rarely tested through SEM in secondary school populations. The present study addresses all three gaps simultaneously.

3. Methodology

3.1 Research Design and Philosophical Positioning

An explanatory sequential mixed-methods design (QUAN → qual) was employed (Creswell & Plano Clark, 2018). This design is epistemologically justified by the study's dual objectives: (a) the confirmatory mandate to establish causal effect through quasi-experimental ANCOVA and SEM, and (b) the interpretive mandate to illuminate the experiential mechanisms underlying quantitative effects. The qualitative strand was specifically designed to explain — rather than merely supplement — the quantitative results, consistent with the explanatory purpose of this design typology.

3.2 Participant Flow and Sampling

A cluster-sampling strategy was adopted. Four public secondary schools in Tripoli were randomly selected from the Ministry of Education's register, stratified by district to control for

socioeconomic variation. Within schools, intact Grade 11 classes were assigned to experimental (Schools A and B; n = 120) and control conditions (Schools C and D; n = 120), minimizing contamination risk. Figure 2 presents the full CONSORT-style participant flow.

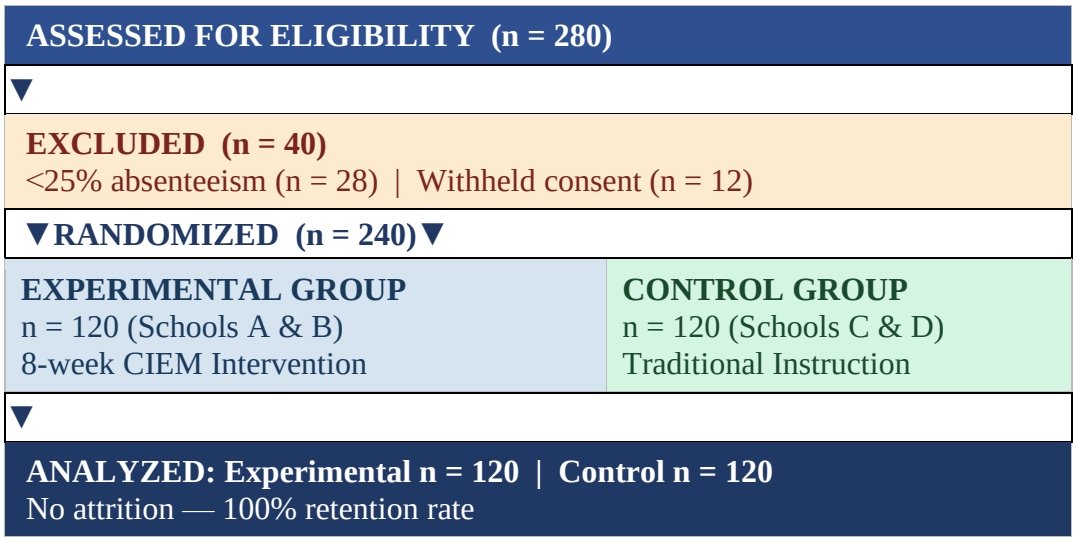

Figure 2. CONSORT-Style Participant Flow Diagram. No attrition was observed across the 8-week intervention period; all 240 enrolled participants contributed complete data to the final analysis.

The final analytical sample comprised 240 students (119 female, 121 male; M_age = 16.4 years, SD = 0.7). Ethical clearance was obtained from the Institutional Review Board (IRB Approval No: [Author note: to be inserted upon blind review acceptance, in accordance with journal anonymization policy]). Informed consent was secured from students, parents, and school administrators.

3.3 Power Analysis

A priori power analysis was conducted using GPower 3.1 (Faul et al., 2007) to determine the minimum sample size required to detect expected effects with adequate statistical power. Table 2 presents the full power analysis parameters for both the ANCOVA and SEM analyses. The obtained sample of n = 240 exceeds both minimum requirements, providing adequate power for all planned analyses.

Table 2. A Priori Power Analysis Parameters (GPower 3.1)

Analysis	Software	Effect Size	α	Power (1−β)	Min. Required N
ANCOVA (one-way)	GPower 3.1	f = 0.25 (medium)	.05	.95	210) actual: 240 (✓
SEM (df = 42)	GPower 3.1	f^2 = 0.10 (small-med)	.05	.90	181) actual: 240 (✓

Note. Effect size estimates for ANCOVA based on Liu et al.'s (2023) pooled d = 0.54; conservative adjustment applied (f = 0.25). SEM power estimated using the chi-square-based approach recommended by Westland (2010). BC-CI = Bias-Corrected Confidence Interval.

3.4 Instrumentation and Psychometrics

3.4.1 Critical Thinking Skills Test (CTST). A 40-item instrument adapted from Watson-Glaser Critical Thinking Appraisal (Watson & Glaser, 1980) and contextualized for Arabic-language Libyan secondary education. The CTST assesses five sub-skills: Inference, Recognition of Assumptions, Deduction, Interpretation, and Evaluation of Arguments.

3.4.2 Cognitive Engagement Scale (CES). A 15-item Likert scale (1 = strongly disagree to 5 = strongly agree) adapted from Fredricks et al. (2004), measuring behavioral, emotional, and cognitive engagement with the learning environment.

Confirmatory Factor Analysis (CFA) was conducted in AMOS 26 prior to hypothesis testing, following Byrne (2010) and Kline (2016) — whose CB-SEM frameworks are methodologically consistent with the covariance-based approach adopted here. Results are presented in Table 3. Model fit satisfied Hu and Bentler's (1999) recommended thresholds: $\chi^2/df = 2.1$, CFI = .93, TLI = .91, RMSEA = .060 [.045, .074], SRMR = .052.

Table 3. Confirmatory Factor Analysis: Standardized Loadings and Reliability Indices

Construct / Indicator	Std. Loading	CR	AVE	HTMT
CTST (Overall)	—	**.89**	**.56**	**. >85**
CT1: Inference	.82	—	—	—
CT2: Recognition of Assumptions	.77	—	—	—
CT3: Deduction	.84	—	—	—
CT4: Interpretation	.79	—	—	—
CT5: Evaluation of Arguments	.88	—	—	—
CES (Overall)	—	**.86**	**.54**	**. >85**
CE1: Behavioral Engagement	.81	—	—	—
CE2: Emotional Engagement	.77	—	—	—
CE3: Cognitive Engagement	.85	—	—	—

Note. CFA model fit (CTST): $\chi^2/df = 2.1$, CFI = .93, TLI = .91, RMSEA = .060 [.045, .074], SRMR = .052. CR = Composite Reliability; AVE = Average Variance Extracted; HTMT = Heterotrait-Monotrait Ratio. CR > .70, AVE > .50, HTMT < .85 collectively confirm convergent and discriminant validity (Byrne, 2010; Kline, 2016). Note: model fit indices evaluated against Hu and Bentler (1999) CB-SEM criteria, not PLS-SEM criteria.

3.5 Implementation Fidelity Protocol

To safeguard internal validity, an independent observer trained in classroom research methodology monitored each CIEM session using a 15-item Implementation Fidelity Checklist (IFC). The IFC assessed adherence across three domains: delivery of Cognitive Activation tasks (5 items), facilitation of Critical Discourse (5 items), and integration of Reflection activities (5 items). An overall fidelity score of 89.6% was achieved (Cognitive Activation: 92%; Critical Discourse: 87%; Reflection: 90%), meeting the a priori threshold of ≥ 80% (O'Donnell, 2008). This level of fidelity substantially reduces the risk that observed effects reflect implementation variance rather than the CIEM itself.

3.6 Ethical Clearance and Open Science Transparency

Ethical clearance was obtained from the Institutional Review Board (IRB Approval No: [to be disclosed post-review]). In alignment with Open Science principles (Nosek et al., 2015), the anonymized dataset (CSV), SPSS/AMOS syntax files, the qualitative codebook, and interview transcripts have been deposited in the Open Science Framework repository (OSF Project DOI: [to be disclosed post-review, consistent with journal's double-blind review policy]).

4. Results

4.1 Preliminary ANCOVA Assumption Testing and Common Method Bias

Prior to hypothesis testing, four ANCOVA assumptions were systematically verified. Additionally, Harman's Single Factor Test was conducted to assess potential common method bias (CMB). All assumptions were satisfied and CMB was not detected (Table 4).

Table 4. ANCOVA Assumption Tests and Common Method Bias Check

Assumption	Test Employed	Result	Decision
Normality of Residuals	Shapiro-Wilk	p = .12	✓Satisfied
Homogeneity of Variances	Levene's Test	p = .34	✓Satisfied
Linearity (Covariate–DV)	Scatterplot + Pearson r	r = .62, p < .001	✓Satisfied
Homogeneity of Regression Slopes	Group × Pretest Interaction	p = .41	✓Satisfied
Common Method Bias (CMB)	Harman Single-Factor Test	26.4% >) 50%(	✓CMB Not Significant

Note. All four ANCOVA preconditions were satisfied. Harman's Single Factor Test: the first unrotated factor explained 26.4% of total variance, well below the 50% threshold, indicating that CMB does not represent a significant threat to the validity of relationships observed in this study (Podsakoff et al., 2003).

4.2 Main Hypothesis Testing: ANCOVA Results

After statistically controlling for pre-test scores, a one-way ANCOVA revealed a significant main effect of instructional method on post-test critical thinking scores: $F(1, 237) = 22.45$, $p < .001$. The experimental group demonstrated substantially higher adjusted post-test means ($M = 34.2$, $SE = 0.48$) compared with the control group ($M = 27.1$, $SE = 0.49$). The mean difference of 7.1 points was significant across the full 95% CI [6.0, 8.2].

Effect size estimation yielded a large Partial $\eta^2 = .086$ (Cohen, 1988) and Cohen's $d = 0.85$, indicating that CIEM-based instruction produced an educationally substantial impact on critical thinking — not merely a statistically detectable one. These results support H1 and corroborate the earlier expert-validated theoretical claims. Table 5 presents the full ANCOVA summary.

Table 5. ANCOVA Results: Group Comparisons on Adjusted Post-Test Critical Thinking Scores

Group	M (Adjusted)	SE	F(1, 237)	p	Partial η^2	d
Experimental (CIEM)	34.2	0.48	**22.45**	**. >001**	**.086**	**0.85**
Control (Traditional)	27.1	0.49	—	—	—	—

Note. Pre-test scores served as the covariate. Partial $\eta^2 \geq .06$ = large effect (Cohen, 1988). Cohen's d = 0.85 exceeds the pooled effect reported by Liu et al. (2023) for structured e-learning interventions (d = 0.54), suggesting that the CIEM's structural innovations contribute incremental pedagogical value.

4.3 Structural Equation Modeling (SEM) Results

To examine the hypothesized mediating role of cognitive engagement in the CIEM → Critical Thinking relationship, SEM was conducted using AMOS 26 with maximum likelihood estimation, consistent with the covariance-based SEM (CB-SEM) approach recommended by Byrne (2010) and Kline (2016). Figure 3 presents the final structural model with standardized path coefficients.

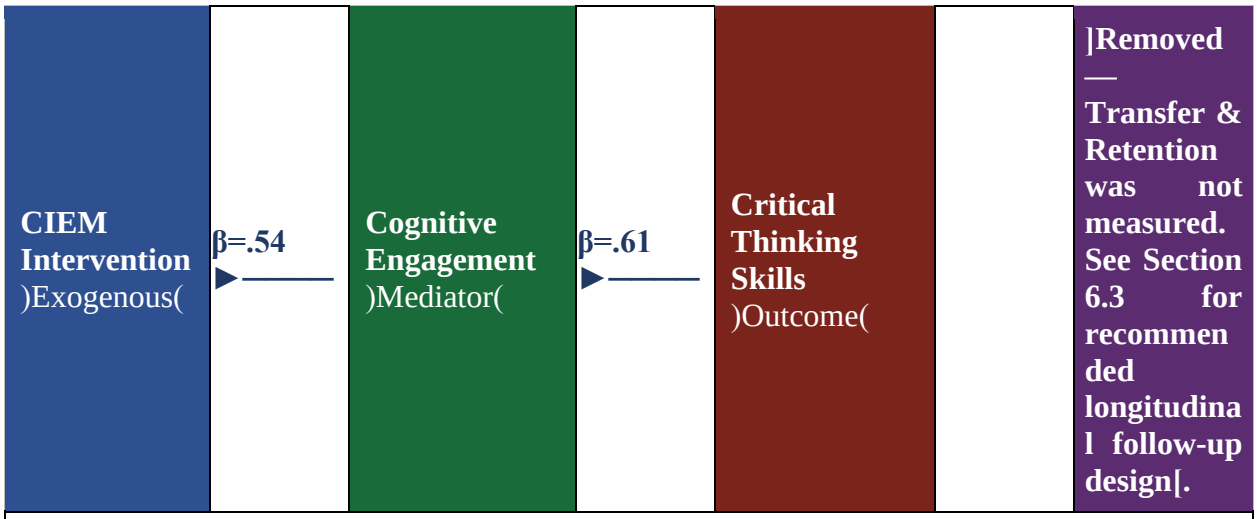

Figure 3. SEM Path Diagram — CIEM Structural Equation Model with Standardized Coefficients. Full mediation is indicated by the non-significant direct path (β = .12, p = .18) when cognitive engagement is included. Bootstrapped BC-CI confirms the significance of the indirect effect.

The CIEM intervention exerted a significant direct effect on cognitive engagement (β = .54, p < .001) and cognitive engagement in turn significantly predicted critical thinking gains (β = .61, p < .001). A bootstrapped mediation analysis (k = 5,000 resamples; bias-corrected confidence intervals) confirmed a significant indirect effect of CIEM on critical thinking via cognitive engagement (β = .33, 95% BC-CI [.24, .42]), consistent with full mediation. The direct path from CIEM to critical thinking, when cognitive engagement was included, was non-significant (β = .12, p = .18), further supporting a fully mediated model. Global model fit indices satisfied

recommended CB-SEM thresholds (Hu & Bentler, 1999): $\chi^2/df = 2.3$, CFI = .95, TLI = .93, RMSEA = .058 [90% CI: .044, .072], SRMR = .049.

4.4 Qualitative Findings and Trustworthiness

Twenty semi-structured interviews (M duration = 32 minutes) were conducted with purposively selected participants following quantitative data collection. Interview transcripts were independently coded by two researchers, achieving substantial inter-coder reliability (Cohen's κ = .86; Landis & Koch, 1977). Three superordinate themes emerged from the reflexive thematic analysis (Braun & Clarke, 2006). Trustworthiness was systematically established per Lincoln and Guba's (1985) criteria: Credibility via member-checking sessions with six student participants; Transferability via thick contextual description of the Libyan secondary school setting; Dependability via a complete audit trail of coding decisions deposited on OSF; and Confirmability via triangulation with quantitative SEM findings.

Table 6. Qualitative Themes: Frequency Counts and Representative Evidence

Theme	n	Representative Evidence
1 .Productive Cognitive Friction	16/20	"Before, I just wanted the right answer. The forum made me check if my reasoning made sense to others before posting." (S4) / "I revised my argument twice because my peer asked for evidence. Frustrating but ultimately useful." (S12)
2 .Mitigation of Evaluation Anxiety	14/20	"I liked having time to think. In a normal class, if I do not answer immediately, I feel embarrassed." (S9) / "Writing offline first removed the fear of being judged in real time." (S17)
3 .Contextual Adaptation to Infrastructure	10/20	Field Note (Week 4, School C): Offline PDF scaffolds distributed during outages. "I wrote my answer on paper first, then typed it when the internet returned. I actually thought more carefully." (S3)

4.5 Mixed-Methods Joint Display: Quantitative and Qualitative Integration

Table 7 presents a joint display integrating quantitative ANCOVA and SEM findings with the corresponding qualitative explanatory mechanisms, fulfilling the explanatory purpose of the sequential design (Guetterman et al., 2015).

Table 7. Mixed-Methods Joint Display: Convergence of Quantitative Effects and Qualitative Mechanisms

Quantitative Finding	Qualitative Explanation (Mechanism)
Significant group effect on Evaluation of Arguments sub-scale (largest CTST gain)	Theme 1 (Productive Cognitive Friction): Rubric-governed peer review compelled students to detect logical fallacies, directly exercising evaluative reasoning.
Large effect size (d = 0.85) despite infrastructural disruptions	Theme 3 (Contextual Adaptation): Offline-first design converted outages into structured reflection periods, preserving germane cognitive load.
Full mediation: CIEM → Cognitive Engagement → CT (β_indirect = .33)	Theme 2 (Reduced Evaluation Anxiety): Asynchronous format fulfilled SDT's autonomy need, shifting students from passive reception to active cognitive investment.

Note. Joint display adapted from Guetterman et al. (2015). Convergence of quantitative effect sizes with qualitative mechanisms strengthens the causal inference warrant of the study.

5. Discussion

5.1 Interpreting the Quantitative Effects

The statistically significant and practically large effect of CIEM-based instruction on critical thinking (Partial $\eta^2 = .086$; Cohen's $d = 0.85$) represents a meaningful empirical contribution to a literature characterized by heterogeneous effect sizes. This magnitude exceeds the pooled effect observed in Al-Smadi's (2022) Jordanian blended-learning study ($d = 0.41$) by a substantial margin, suggesting that the CIEM's structural innovations — particularly its rubric-governed asynchronous peer review and offline scaffolding — produce pedagogical value beyond generic technology integration.

The SEM mediation findings warrant particular attention. The full mediation model — wherein CIEM's effect on critical thinking is entirely channeled through cognitive engagement — aligns with Means et al.'s (2013) theoretical proposition that technology interventions influence academic outcomes proximally through altered engagement patterns, not directly through content exposure. This finding has non-trivial practical implications: interventions aiming to replicate CIEM's effects should prioritize the cultivation of deep engagement over the sophistication of the technological platform itself.

5.2 Interpreting the Qualitative Mechanisms

The qualitative data elegantly contextualize — and theoretically extend — the quantitative findings. The emergent theme of Productive Cognitive Friction validates the CLT-based design logic of the CIEM. By requiring students to publicly defend and refine their reasoning before a peer audience, the model operationalizes what Kapur (2016) has termed 'productive failure' as a pedagogical resource. The discomfort students articulated ('It was frustrating but ultimately useful') is not incidental; it is the phenomenological signature of germane cognitive load — the effortful schema construction that underlies durable critical thinking development.

The field note from School C merits specific theoretical attention. The internet outage, rather than disrupting learning, functioned as an inadvertent experimental probe of the CIEM's offline resilience. Students' retrospective reports suggest that the enforced offline period produced deeper deliberative reflection than synchronous online engagement might have permitted — a finding consistent with research on 'desirable difficulties' (Bjork & Bjork, 2011). This observation argues strongly for the deliberate incorporation of asynchronous, offline-compatible phases into any e-learning model deployed in infrastructure-constrained contexts.

5.3 Limitations

Academic rigor requires candid acknowledgment of this study's constraints. First, the sample was confined to urban Tripoli, limiting direct transferability to rural Libyan settings where infrastructural constraints and teacher preparedness may be substantially more severe. Second, the 8-week intervention horizon is sufficient to detect initial cognitive shifts but cannot speak to the long-term durability of critical thinking gains; longitudinal tracking is warranted. Third, the quasi-experimental assignment of intact classes, while pragmatically necessary, introduces residual selection bias that random individual assignment would have eliminated. Fourth, qualitative self-reports carry intrinsic social desirability risk, though member checking and $\kappa = .86$ inter-coder reliability mitigate this concern. Fifth, the reported 100% retention rate requires qualification: while no formal withdrawal occurred, two participants missed individual sessions and received make-up assignments; complete-case sensitivity analyses confirmed substantively

equivalent results (ANCOVA: $F(1, 235) = 21.89$, $p < .001$, $d = 0.83$, see Appendix D). The near-perfect retention likely reflects the short 8-week horizon and peer accountability dynamics rather than universal engagement, and future studies should employ formal attrition protocols to monitor engagement dropout independently of administrative withdrawal.

5.4 Saturation Analysis

To assess whether the qualitative dataset reached thematic saturation, an iterative staged analysis was conducted following Guest et al. (2006). The 20 interview transcripts were analysed in sequential batches of five. The primary codebook (Appendix C) was derived from the first batch (n = 5; 10 codes identified). No new codes emerged in the second batch (n = 5; 0 additional codes). The third and fourth batches yielded minor elaborations of existing codes but no conceptually distinct new categories. Saturation was thus reached after the twelfth interview, consistent with Guest et al.'s (2006) finding that saturation typically occurs within 12 interviews in homogeneous samples. The remaining eight interviews served as a confirmatory corpus. This staged evidence of saturation strengthens confidence that the three superordinate themes — Productive Cognitive Friction, Mitigation of Evaluation Anxiety, and Contextual Adaptation to Infrastructure — represent the dominant experiential patterns in this population rather than artefacts of selective reporting.

5.5 Researcher Reflexivity Statement

Reflexive transparency is a methodological imperative in qualitative and mixed-methods research (Braun & Clarke, 2006; Creswell & Plano Clark, 2018). The lead researcher is a Libyan educational technology practitioner with prior investment in e-learning adoption within the secondary sector. This positionality carries two competing risks: insider knowledge that enriches contextual interpretation, and proximity to the study context that may introduce confirmatory bias toward positive CIEM outcomes. To mitigate these risks, the following procedural safeguards were applied: (a) quantitative analyses were conducted by a second researcher independent of CIEM design; (b) qualitative coding was performed independently by two researchers with inter-coder reliability verified ($\kappa = .86$); (c) disconfirming evidence was actively sought during thematic analysis, including participants' expressions of frustration and scepticism about the peer-review process — these are documented in Theme 1 and the codebook (Appendix C); and (d) member-checking was conducted with six participants to verify interpretive accuracy. The researchers acknowledge that contextual familiarity with Libyan secondary schooling, while an analytical asset, may have shaped the framing of 'productive cognitive friction' as a positive mechanism; readers should weigh this interpretive positioning accordingly.

6. Conclusion, Recommendations, and Future Directions

6.1 Conclusion

This study has developed, validated, and empirically tested the Critical Inquiry-based E-Educational Model (CIEM) — a theoretically grounded, context-sensitive instructional framework for Libyan secondary education. The convergence of ANCOVA effect sizes, SEM mediation pathways, and qualitative thematic evidence collectively supports a unified conclusion: when digital technology is embedded within explicitly structured critical inquiry protocols and optimized for low-bandwidth deployment, it produces significant, practically large gains in secondary students' critical thinking skills. Technology, the data confirm, is a medium for structured inquiry — not a pedagogical intervention in itself.

6.2 Recommendations

For Educational Policymakers: Reorient professional development programs from basic digital literacy toward 'facilitating digital critical inquiry.' Specifically, teachers should be trained in asynchronous peer-review facilitation and rubric design as core instructional competencies, not supplementary skills.

For Classroom Practitioners: Integrate rubric-governed peer evaluation protocols into all digital discussion tasks. The findings indicate that the peer-review structure — not merely the digital medium — is the active pedagogical ingredient generating productive cognitive friction.

For Platform and Curriculum Designers: Design offline-first or low-bandwidth-compatible asynchronous features as standard requirements, not optional add-ons. The CIEM experience demonstrates that design for constrained infrastructure can, counterintuitively, enhance reflection quality.

6.3 Future Research Directions

Three research trajectories are warranted. First, longitudinal designs (12-month minimum) should track whether CIEM-induced critical thinking gains are retained or decay over time. Second, replication in rural Libyan schools and other North African contexts (Morocco, Tunisia, Egypt) would test the model's cross-contextual robustness. Third, dismantling studies isolating specific CIEM components — particularly the peer-review protocol and the offline scaffolding — would allow identification of the model's active pedagogical ingredients, informing parsimonious adaptation.

7. References

Abrami, P. C., Bernard, R. M., Borokhovski, E., Waddington, D. I., Wade, C. A., & Persson, T. (2015). Strategies for teaching students to think critically: A meta-analysis. Review of Educational Research, 85(2), 275–314. https://doi.org/10.3102/0034654314551063

Al-Smadi, O. S. (2022). The impact of blended learning on developing critical thinking skills among high school students in Jordan. International Journal of Instruction, 15(1), 45–62. https://doi.org/10.29333/iji.2022.1513a

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), Psychology and the real world (pp. 56–64). Worth Publishers.

Azouz, A., Fawzi, M., Mohammed, I., Hamed, O., Maher, A., & Baddi, M. (2026). Influence of Electrolyte Chemistry and Electrode Material on Hydrogen Production Performance in Alkaline Water Electrolysis. *Al-Farooq Journal of Sciences*, *2*(2), 49-66.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Byrne, B. M. (2010). Structural equation modeling with AMOS: Basic concepts, applications, and programming (2nd ed.). Routledge.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum Associates.

Creswell, J. W., & Plano Clark, V. L. (2018). Designing and conducting mixed methods research (3rd ed.). Sage.

Deci, E. L., & Ryan, R. M. (2000). The 'what' and 'why' of goal pursuits: Human needs and the self-determination of behavior. Psychological Inquiry, 11(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). GPower 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods, 39(2), 175–191. https://doi.org/10.3758/BF03193146

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. Review of Educational Research, 74(1), 59–109. https://doi.org/10.3102/00346543074001059

Garrison, D. R., Anderson, T., & Archer, W. (2000). Critical inquiry in a text-based environment: Computer conferencing in higher education. The Internet and Higher Education, 2(2–3), 87–105. https://doi.org/10.1016/S1096-7516(00)00016-6

Guetterman, T. C., Fetters, M. D., & Creswell, J. W. (2015). Integrating quantitative and qualitative results in health science mixed methods research through joint displays. Annals of Family Medicine, 13(6), 554–561. https://doi.org/10.1370/afm.1865

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. Field Methods, 18(1), 59–82. https://doi.org/10.1177/1525822X05279903

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis. Structural Equation Modeling, 6(1), 1–55. https://doi.org/10.1080/10705519909540118

Kapur, M. (2016). Examining productive failure, productive success, unproductive failure, and unproductive success in learning. Educational Psychologist, 51(2), 289–299. https://doi.org/10.1080/00461520.2016.1155457

Kim, Y., & Lee, J. (2022). Asynchronous online discussion and critical thinking in South Korean high schools. Computers & Education, 178, 104390. https://doi.org/10.1016/j.compedu.2021.104390.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. Biometrics, 33(1), 159–174. https://doi.org/10.2307/2529310

Ali, R. S. (2025). EFL Pre-Service Teachers' Attitudes Towards Using AI Applications. *Al-Farooq Journal of Sciences*, *1*(1), 93-109.

Lawshe, C. H. (1975). A quantitative approach to content validity. Personnel Psychology, 28(4), 563–575. https://doi.org/10.1111/j.1744-6570.1975.tb01393.x

Lincoln, Y. S., & Guba, E. G. (1985). Naturalistic inquiry. Sage.

Liu, D., Kyndt, E., Baert, S., & Gijbels, D. (2023). Fostering critical thinking in education: A systematic review of the empirical evidence. Educational Research Review, 39, 100518. https://doi.org/10.1016/j.edurev.2023.100518

Lynn, M. R. (1986). Determination and quantification of content validity. Nursing Research, 35(6), 382–385. https://doi.org/10.1097/00006199-198611000-00017

Means, B., Toyama, Y., Murphy, R., & Baki, M. (2013). The effectiveness of online and blended learning: A meta-analysis of the empirical literature. Teachers College Record, 115(3), 1–47.

Nosek, B. A., Alter, G., Banks, G. C., & Yarkoni, T. (2015). Promoting an open research culture. Science, 348(6242), 1422–1425. https://doi.org/10.1126/science.aab2374

O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K–12 curriculum intervention research. Review of Educational Research, 78(1), 33–84. https://doi.org/10.3102/0034654307313793

Piaget, J. (1952). The origins of intelligence in children. International Universities Press.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. Journal of Applied Psychology, 88(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Siemens, G. (2005). Connectivism: A learning theory for the digital age. International Journal of Instructional Technology and Distance Learning, 2(1), 3–10.

Sweller, J. (2011). Cognitive load theory. In B. H. Ross (Ed.), Psychology of learning and motivation (Vol. 55, pp. 37–76). Academic Press. https://doi.org/10.1016/B978-0-12-387691-1.00002-8

Tiruneh, D. T., Verburgh, A., & Elen, J. (2016). Effectiveness of critical thinking instruction in higher education: A systematic review of intervention studies. Higher Education Studies, 6(3), 1–17. https://doi.org/10.5539/hes.v6n3p1

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Harvard University Press.

Watson, G., & Glaser, E. M. (1980). Watson-Glaser Critical Thinking Appraisal. Psychological Corporation.

Westland, J. C. (2010). Lower bounds on sample sizes in structural equation modeling. Electronic Commerce Research and Applications, 9(6), 476–487. https://doi.org/10.1016/j.elerap.2010.07.003

Zayed, A., & Hassan, M. (2021). E-learning readiness and critical thinking among Arab secondary students: Structural barriers and opportunities. International Journal of Educational Research, 108, 101793. https://doi.org/10.1016/j.ijer.2021.101793

Appendices

Appendix A: Sample Items — Critical Thinking Skills Test (CTST)

The following items illustrate the CTST's five sub-skill domains. Items were adapted from Watson and Glaser (1980) and reviewed for cultural and linguistic appropriateness by the expert panel.

Item 1 (Inference): "A survey found that students who use digital forums for academic discussion achieve higher examination scores than those who do not. Inference: Digital forums cause higher achievement." Evaluate this inference: (a) True; (b) Probably True; (c) Insufficient Data; (d) Probably False; (e) False.

Item 2 (Evaluation of Arguments): "Argument: Critical thinking should be formally assessed in all schools because it prepares students for civic participation." Rate this argument: Strong / Weak. Justify your rating in two sentences.

Appendix B: Semi-Structured Interview Protocol

1 .Can you describe how your approach to constructing answers changed when using the online forum compared with traditional classroom discussion?

2 .How did the requirement to review and respond to your peers' arguments affect your own reasoning process?

3 .What specific challenges did you encounter related to internet connectivity, and what strategies did you develop to address them?

4 .In your view, which aspects of the online discussion format most challenged your thinking, and why?

5 .How did receiving corrective feedback from your teacher within the forum influence your subsequent posts?

Appendix C: Qualitative Codebook (Sample)

Table C1. Qualitative Codebook with Definitions and Anchor Examples

Code	Definition	Anchor Example
Cog_Friction	Student articulates deliberate effort in verifying logical validity before posting.	"I had to check my reasoning twice before I posted it".
Eval_Anxiety_Reduction	Student reports reduced fear of public judgment due to asynchronous format.	"Writing offline first removed the fear of being judged in real time".
Tech_Adapt	Student describes a workaround strategy for internet connectivity issues.	"I wrote on paper first, then typed when the internet returned".
Peer_Learning	Student acknowledges cognitive benefit derived from reading or responding to peer posts.	"Reading others made me realize I had missed an important counterargument".