

A Qualitative Thematic Analysis of ChatGPT-4 Integration in Extensive Reading to Enhance Critical Thinking in EFL Higher Education

Fina Khiyarotun Nisa¹, Rahayu Puji Haryanti²

^{1,2}Department of English Education,
Languages and Arts College, Semarang State University, Semarang, Indonesia.

Corresponding Author: Fina Khiyarotun Nisa

DOI: <https://doi.org/10.52403/ijrr.20260815>

ABSTRACT

Integrating Artificial Intelligence (AI), specifically Large Language Models like ChatGPT-4, into English as a Foreign Language (EFL) instruction offers significant pedagogical opportunities, yet its qualitative impact on higher-order cognitive processing remains nuanced. This study provides an in-depth qualitative thematic analysis examining how ChatGPT-4 is integrated into Extensive Reading classes to foster critical thinking skills, evaluating the resulting cognitive transformations, identifying learner challenges, and capturing instructional reflections. Using a convergent parallel mixed-method design grounded in Facione's Critical Thinking Framework and Vygotsky's Social Constructivism, data were collected from 48 second-semester EFL undergraduate students and their instructor at University X, Kudus, Java, using classroom observations, semi-structured interviews, student worksheets, and standardized pre-/post-tests. Thematic analysis of qualitative data revealed five major thematic dimensions: (1) Guided Pedagogical Integration & Scaffolding, (2) Cognitive Deepening in Textual Processing, (3) Collaborative Critical Dialogue & Peer Evaluation, (4) Operational & Epistemic Challenges, and (5) Instructional Reflection on AI Mediation. Quantitative findings

corroborated these thematic insights, demonstrating a statistically meaningful improvement in critical thinking scores (Pre-test $M = 49.7$; Post-test $M = 73.5$; Normalized Gain $g = 0.65$, Moderate). The findings demonstrate that ChatGPT-4 operates effectively as an artificial mediational tool within the Zone of Proximal Development when paired with structured prompts and teacher-led reflection, transforming learners from passive consumers into active, analytical evaluators.

Keywords: ChatGPT-4, Extensive Reading, Critical Thinking, Thematic Analysis, EFL, Artificial Intelligence in education.

INTRODUCTION

The rapid evolution of generative artificial intelligence (AI) technologies, particularly Large Language Models (LLMs) such as OpenAI's ChatGPT series, has fundamentally reshaped the landscape of higher education (1). However, as educators embrace digitalization, a central dilemma emerges: whether AI assistance fosters intellectual autonomy or inadvertently cultivates passive learning habits and cognitive overreliance (2). In foreign language education, reading is not merely a passive decoding process but a complex, interactive cognitive endeavor requiring

textual interpretation, analysis, and critical evaluation(3). Extensive Reading, characterized by the line-by-line detailed examination of short, complex texts focusing on syntax, vocabulary, argument structure, and contextual nuances(4), serves as an ideal pedagogical site for cultivating higher-order cognitive processing. Despite its potential, traditional Extensive Reading instruction often suffers from limited practice time, rigid teacher-selected texts, and lower student engagement, leading to surface-level reading comprehension rather than critical textual critique.

Initial diagnostic assessments conducted in the English Education Department at University X in Kudus, Central Java, revealed notable deficits in students' critical thinking skills when mapped against Facione's framework: Analysis, Evaluation, and Inference(2). While second-semester EFL students demonstrated basic competence in identifying explicit factual information (interpretation), they struggled significantly with analyzing argument structures, identifying implicit assumptions, evaluating source credibility, drawing deductive inferences, and exercising metacognitive self-regulation. Without structured instructional interventions, students frequently succumbed to uncritical acceptance of printed and digital information, including AI outputs(5).

To address these empirical and pedagogical gaps, this study systematically investigates the integration of ChatGPT-4 into Extensive Reading classes as a mediational tool to promote critical thinking. The primary aims of this research paper are to: (1) analyze how ChatGPT-4 is pedagogically integrated into Extensive Reading instruction, (2) evaluate the qualitative themes emerging from students' cognitive development, (3) explore student challenges and perceived benefits, and (4) examine the lecturer's critical reflections on AI-assisted pedagogy.

MATERIALS & METHODS

This study employed a convergent parallel mixed-methods design to capture both the

qualitative thematic nuances and quantitative effectiveness of ChatGPT-4 integration. The research was conducted in the English Education Department at University X, Kudus, Central Java, during the second semester of the 2024/2025 academic year.

Participants & Sampling

Forty-eight (48) undergraduate EFL students enrolled in two Extensive Reading classes participated in the study, alongside their primary course lecturer. Participants were selected by purposive sampling based on their prior exposure to Extensive Reading in semester one and their willingness to engage with AI-supported instructional interventions.

Data Collection Procedures & Instruments

Data were gathered using four complementary instruments: (1) Standardized Critical Thinking Assessment Tests adapted from Assessment Day UK (comprising 86 items spanning Inference, Assumptions, Deduction, Interpretation, and Argumentation), administered pre- and post-intervention; (2) Classroom Observation Sheets with checklist markers and detailed field notes across four 130-minute instructional sessions; (3) In-depth Semi-Structured Interviews with 14 purposively selected students and the course lecturer; and (4) Document Analysis of course syllabi (RPS), lesson plans, student worksheets, and AI interaction logs.

Data Analysis Techniques

Quantitative test scores were analyzed using descriptive statistics and the Normalized Gain (N-Gain) formula: $N\text{-Gain} = (\text{Post_score} - \text{Pre_score}) / (\text{Max_score} - \text{Pre_score})$. Qualitative data from interview transcripts, observation notes, and student worksheets were analyzed strictly using Braun and Clarke's 2006 six-phase thematic analysis framework (Familiarization, Initial Code Generation, Theme Searching, Theme Reviewing, Defining and Naming Themes,

and Report Production). Qualitative Data Analysis (QDA) Miner Lite software was utilized for rigorous text segment coding, category sorting, and thematic mapping.

RESULT

The findings are structured to prioritize the detailed qualitative thematic analysis of ChatGPT-4's integration in Extensive Reading, supported by quantitative pre-/post-test gains.

1. Quantitative Improvement in Critical Thinking Skills

The administration of the Critical Thinking Assessment Test to 48 students demonstrated a substantial improvement in overall cognitive performance following the ChatGPT-4 integrated intervention. As shown in Table 1, the baseline pre-test mean score was 49.7 (SD = 10.4, classified as Fair), whereas the post-test mean score rose to 73.5 (SD = 7.2, classified as Good), representing an absolute mean score gain of 23.8 points.

Table 1: Descriptive Statistics and Normalized Gain (N-Gain) of Critical Thinking Test (N=48)

Assessment Phase	Mean Score	Percentage (%)	Qualitative Performance Category	Mean N-Gain	N-Gain Category (Hake, 2002)
Pre-Test	49.7	57.8%	Fair Critical Thinking Level	0.65	Moderate Improvement
Post-Test	73.5	85.5	Good to Excellent Critical Thinking Level		

The calculated average Normalized Gain score ($g = 0.65$) falls squarely within Hake's (2002) Moderate Improvement range ($0.30 \leq g < 0.70$). This quantitative trajectory confirms that structured pedagogical integration of ChatGPT-4 significantly enhances students' capacity to make inferences, evaluate premises, recognize unstated assumptions, and deduce valid conclusions.

2. Qualitative Thematic Analysis of ChatGPT-4 Integration

Rigorous thematic coding of interview transcripts, observation records, and student work samples yielded five core thematic dimensions and ten underlying sub-themes, mapping the cognitive and pedagogical dynamics of AI integration in Extensive Reading.

Table 2: Comprehensive Qualitative Thematic Analysis Matrix

Core Theme	Sub-Themes	Key Structural Indicators & Observed Behaviors	Facione's CT Dimension
Theme 1: Scaffolding & Guided AI Integration	Vocabulary & Structural Clarification	Querying contextual word meanings, breaking down technical idioms (e.g., "Ring of Fire"), verifying complex syntax.	Interpretation
	Guided Prompts & Worksheet Alignment	Utilizing teacher-designed worksheets to post targeted analytical prompts rather than seeking direct answers.	Interpretation/Analysis
Theme 2: Cognitive Deepening & Textual Evaluation	Comparative Argument Critique	Contrasting ChatGPT's alternative perspectives against author arguments; identifying weak supporting evidence.	Analysis / Evaluation
	Assumption Uncovering & Deduction	Detecting underlying implicit biases, testing logical fallacies, and verifying AI claim credibility against texts.	Evaluation / Inference
Theme 3: Collaborative Dialogue & Metacognition	Peer Cross-Verification	Group debate on AI outputs, negotiating meaning, defending personal interpretations against peer and AI claims	Explanation
	Self-Regulation & Reflection	Re-evaluating initial personal assumptions, revising textual	Self-Regulation

		interpretations upon finding AI and text discrepancies.	
Theme 4: Operational & Epistemic Challenges	Information Validation	Encountering AI fact fabrications, lack of academic citation, over-generalized or superficial commentary.	Evaluation
	Overreliance Risk	Resisting copy-paste temptations, initial cognitive friction when prompted to critique AI responses.	Self-Regulation
Theme 5: Instructional & Pedagogical Reflection	Role Shift from Authority to Facilitator	Lecturer guiding critical inquiry rather than transmitting answers; structuring ethical boundaries.	Pedagogical Alignment
	Balanced AI Integration	Pairing AI individual processing with offline human centered reflective class discussions.	Constructivist Mediation

Theme 1: Scaffolding and Guided AI Integration

The first theme highlights how ChatGPT-4 functions as an immediate linguistic and structural scaffold during the initial reading phase. Rather than replacing reading, ChatGPT-4 resolved lower-level linguistic barriers (e.g., obscure idioms, complex passive structures, specialized vocabulary), enabling students to allocate higher cognitive resources toward critical analysis.

"I used ChatGPT to ask about difficult vocabulary and complex phrases like 'held captive' or 'Ring of Fire'. When I understand the exact context of the words instantly, I can focus on whether the author's overall message makes sense." — [Student 3, Interview]

Crucially, observation data confirmed that integration succeeded because the lecturer enforced a "guided prompt strategy" via structured worksheets. Students were prohibited from prompting "Give me the answer," but were trained to ask structured queries: *"What is the author's argument of this text?"* followed by *"In what ways does your interpretation differ from my group's understanding?"*

Theme 2: Cognitive Deepening and Textual Evaluation

The second theme captures the core transition from passive comprehension to active critical evaluation. When prompted to compare ChatGPT-4's generated counter-arguments with the original text, students performed higher-level cognitive operations

in the Analysis, Evaluation, and Inference dimensions.

"When we asked ChatGPT to provide a contrasting argument regarding social media's impact, ChatGPT gave a very general answer. But when we checked the author's original text, the author provided specific statistical evidence. So my group decided the author's argument was much stronger than ChatGPT's response." — [Student 11, Group Worksheet & Interview]

This comparative exercise forced students to evaluate premise strength, spot unstated assumptions, and judge argument validity. Rather than viewing AI as an infallible oracle, students learned to treat ChatGPT-4 as an unverified interlocutor whose claims required rigorous textual verification.

Theme 3: Collaborative Dialogue and Metacognitive Self-Regulation

The third theme underscores the social constructivist dimension of the classroom. After individual interactions with ChatGPT-4, students entered peer-group discussions in which AI-generated outputs were debated. This phase catalyzed the Explanation and Self-Regulation dimensions.

"In my group, two members agreed with ChatGPT's summary, but I noticed ChatGPT missed a crucial nuance in paragraph 4. We read the paragraph line by line again, and realized ChatGPT had oversimplified the conclusion. That made me realize I need to double-check my own

quick assumptions too." — [Student 8, Interview]

Peer collaboration prevented isolated overreliance on AI and stimulated metacognitive reflection, as students actively defended their positions and corrected their analytical errors through dialogue.

Theme 4: Operational and Epistemic Challenges

The fourth theme delineates the operational obstacles encountered by learners. Students identified key AI limitations, including generic superficial commentary and lack of rigorous academic sources. Furthermore, students noted the constant psychological temptation to rely on automated outputs.

"Sometimes ChatGPT sounds very confident and convincing, but when you look closely at the reading text, the AI completely invented details that weren't there." — [Student 14, Interview]

Navigating these challenges cultivated a crucial digital literacy skill: epistemic skepticism. Learners recognized that AI efficiency must be balanced with critical vigilance.

Theme 5: Instructional Reflection on AI Mediation

The final theme reflects the course lecturer's pedagogical insights regarding AI integration. The lecturer noted that successfully integrating ChatGPT-4 required a shift from direct instruction to task design and cognitive facilitation.

"ChatGPT forced me to think about my teaching role. If I just ask basic comprehension questions, AI renders the homework meaningless. But when I design tasks where students must critique AI-generated logic against the text, ChatGPT becomes an incredible cognitive exercise. The key is maintaining a strict approach where offline peer discussion covers the learning." — [Course Lecturer, Interview]

DISCUSSION

The quantitative gain in students' critical thinking scores (N-Gain=0.65) substantiates the cognitive benefits of conversational AI in language learning. From a critical pedagogical standpoint, this improvement cannot be attributed merely to the automated capabilities of ChatGPT-4, but rather to its structured function as an artificial mediational tool within the learner's Zone of Proximal Development. This allows students to reallocate processing capacity toward higher-order dimensions of Analysis, Evaluation, and Inference. The success of this cognitive reallocation depends strictly on worksheet-guided prompt design; without explicit instructional constraints prohibiting direct answer retrieval, AI mediation risks degenerating into cognitive passivity.

The qualitative and quantitative findings of this study provide critical theoretical and empirical insights into the integration of Generative AI in language education. The observed increase in critical thinking scores (N-Gain = 0.65) directly aligns with prior investigations(6), which asserted that conversational AI can significantly elevate higher-order cognitive engagement when paired with active learning methodologies.

Theoretically, the qualitative findings validate Social Constructivism and Mediational Tool Theory within a modern digital paradigm. When structured through guided worksheets, ChatGPT-4 functions effectively as an artificial mediational tool within the learner's Zone of Proximal Development (ZPD). By providing immediate scaffolding—such as contextualizing complex syntax or generating alternative perspectives—ChatGPT-4 reduces extraneous cognitive load(7), enabling learners to focus their conscious cognitive processing on higher-tier dimensions: Analysis, Evaluation, and Inference. Furthermore, the findings directly address the grave concerns raised regarding AI-induced cognitive passivity (8).

The thematic analysis reveals that overreliance and superficial engagement

occur primarily when AI is utilized in an unguided, unmonitored environment. When pedagogical design incorporates mandatory counter-evaluation, requiring students to explicitly critique AI outputs against authoritative textual evidence, the potential pitfall of AI is transformed into a valuable pedagogical asset. Spotting AI errors and superficial generalizations actively sharpens students' evaluative skills and fosters epistemic skepticism. Finally, the lecturer's reflection highlights the critical importance of maintaining a "human-in-the-loop" pedagogical design(9). ChatGPT-4 serves as an effective cognitive gym during individual processing, but deeper conceptual consolidation, moral reasoning, and metacognitive self-regulation are anchored through offline, peer-to-peer dialogue and teacher-guided reflection. Thus, AI does not replace human instructional agency; rather, it elevates the necessity for sophisticated, constructivist task design in higher education.

CONCLUSION

This study demonstrates that integrating ChatGPT-4 into EFL Extensive Reading instruction effectively enhances higher order critical thinking skills when guided by constructivist principles. Qualitative thematic analysis reveals that ChatGPT-4 acts as a powerful mediational scaffold, facilitating vocabulary clarification, argument critique, assumption identification, and metacognitive self-regulation. Quantitative results corroborate these thematic findings, showing a moderate and statistically meaningful score gain ($N\text{-Gain} = 0.65$). To prevent cognitive passivity and overreliance, AI tools must be embedded within structured instructional frameworks that require students to critically evaluate, cross-verify, and debate AI outputs against textual evidence.

Declaration by Authors

Acknowledgement: None

Source of Funding: None

Conflict of Interest: No conflicts of interest declared.

REFERENCES

1. Yeh HC. The synergy of generative AI and inquiry-based learning: transforming the landscape of English teaching and learning. *Interact Learn Environ.* 2025;33(1):88–102. Available from: <https://doi.org/10.1080/10494820.2024.2335491>
2. Faizin HA, Sutantohadi A, Noviabahari JL, Permatasari I. Generative Artificial Intelligence in the EFL Context: Insights from the Higher Education Sector. *Educational Dynamics: International Journal of Education and Social Sciences.* 2025; 2(2); 70-76. DOI:10.70062/educational dynamics. v1i3.170
3. Renandya WA, Hamied FA, Nurkamto J. The Journal of Asia TEFL English Language Proficiency in Indonesia: Issues and Prospects 1). *J Asia Tefl.* 2018;15(3):618–29. Available from: <http://dx.doi.org/10.18823/asiatefl.2018.15.3.4.6>
4. Eftekhary AA, Kalayeh KB. The relationship between critical thinking and extensive reading on Iranian intermediate EFL learners. *J Nov Appl Sci.* 2014;3(6):623–8.
5. Ferdousi B. Generative Artificial Intelligence ChatGPT in Education: Challenges and Opportunities. 2024;22(6):48–64.
6. Jiang X, Chen C-Hung. Enhancing Critical Thinking Skills with ChatGPT-Powered Activities in Chinese Language Classrooms. 2024; 5:47–73. <https://doi.org/10.46451/ijclt.20240105>
7. Downes S. *Connectivism and Connective Knowledge Essays on meaning and learning networks.* ISBN: 978-1-105-77846-9
8. Chiu TKF, Ahmad Z, Ismailov M, Temitayo I. What are artificial intelligence literacy and competency? A comprehensive framework to support them. *Comput Educ Open [Internet].* 2024;6(March):100171. Available from: <https://doi.org/10.1016/j.caeo.2024.100171>
9. Samala AD, Zhai X, Aoki K, Bojic L, Zikic S. An In-Depth Review of ChatGPT's Pros and Cons for Learning and Teaching in Education. *Int J Interact Mob Technol.* 2024;18(2):96–117.

How to cite this article: Fina Khiyarotun Nisa, Rahayu Puji Haryanti. A qualitative thematic analysis of ChatGPT-4 integration in extensive reading to enhance critical thinking in EFL higher education. *International Journal of Research and Review.* 2026; 13(8): 137-142. DOI: <https://doi.org/10.52403/ijrr.20260815>
