



Critical Thinking in a GenAI World: Evaluating the Efficacy of ChatGPT-Assisted Essay Writing in ESL Classrooms

Dr. E. Ram Bhaskar Raju, E. Namratha

¹Assistant Professor of English, Department of English, Kakatiya Govt. College (Autonomous), Hanumakonda, Telangana, India
E-mail: rambhaskarraju@gmail.com

ORCID: 0000-0003-4287-8914

²Ph.D Research Scholar, Department of English, Faculty of Arts, University of Delhi, India

E-mail: namrathaerra@gmail.com

ORCID: 0009-0005-4625-7762

Received: 29 Nov 2022; Received in revised form: 23 Dec 2022; Accepted: 27 Dec 2022; Available online: 31 Dec 2022

©2022 The Author(s). Published by Infogain Publication. This is an open access article under the CC BY license

(<https://creativecommons.org/licenses/by/4.0/>).

Abstract— The brisk integration of Generative Artificial Intelligence (GenAI), particularly Large Language Models (LLMs) like ChatGPT, has fundamentally disrupted educational paradigms, specifically second language acquisition (SLA) and English as a Second Language (ESL) writing instruction. While GenAI tools offer unparalleled scaffolding for grammar, vocabulary generation, and structural organisation, their impact on higher-order cognitive skills remains highly contested. This study evaluates the empirical efficacy of ChatGPT-assisted essay writing in university-level ESL classrooms, focusing specifically on its effects on students' critical thinking (CT) skills. Using a mixed-methods convergent parallel design over a 15-week academic semester, this study examined $N = 124$ undergraduate ESL students split into an experimental group (utilising ChatGPT as a structured co-writing and dialogue partner) and a control group (utilising conventional process-writing methodologies). Quantitative data were gathered using pre- and post-test assessments scored via a validated Critical Thinking Rubric (evaluating argument analysis, evidence evaluation, alternative perspectives, and bias recognition) along with automated linguistic analysis of student essays. Qualitative data were obtained via semi-structured focus group interviews and student reflective journals. The quantitative findings indicate that while the experimental group demonstrated significant improvements in surface-level linguistic accuracy, lexical density, and structural cohesion, their independent scores in evidence evaluation and independent thesis formulation were statistically lower than those of the control group when ChatGPT assistance was removed in unassisted proctored post-tests ($p < .05$). The qualitative thematic analysis revealed a phenomenon termed "cognitive outsourcing," where students routinely accepted AI-generated assertions without epistemological verification. However, a sub-cohort trained in aggressive prompt engineering and dialectical inquiry demonstrated elevated critical engagement. This paper proposes a "Dialectical AI-Scaffolding Framework" designed to shift ChatGPT from a product-generation tool to an intellectual sparring partner, ensuring that GenAI advances rather than compromises critical thinking in ESL pedagogy.

Keywords— Generative AI, ChatGPT, ESL Writing, Critical Thinking, Cognitive Outsourcing, Second Language Acquisition, Pedagogy.

I. INTRODUCTION

The landscape of English as a Second Language (ESL) writing instruction has historically evolved through major paradigm shifts, transitioning from the behaviourist product-oriented approach to the cognitive process-oriented approach, and eventually to socio-cultural genre-based frameworks. Each shift redefined the locus of agency in textual production. However, the commercial release of OpenAI's ChatGPT and subsequent Generative Artificial Intelligence (GenAI) systems represents an unprecedented, disruptive inflexion point.

For ESL learners, the challenges of academic writing are dual-layered: they must simultaneously manage lower-order linguistic mechanics (morph syntax, lexical retrieval, orthography) and higher-order rhetorical demands (argumentation, logical flow, evidence evaluation, and critical positioning). Historically, the cognitive load imposed by lower-order mechanical constraints often bottlenecked an ESL student's capacity to engage deeply with higher-order critical thinking. GenAI effectively removes this linguistic bottleneck by acting as an instantaneous, hyper-competent mechanical editor and structural scaffolding agent.

However, this technological liberation introduces a profound pedagogical paradox. If an AI agent can instantaneously synthesise literature, resolve grammatical incongruities, balance rhetorical structures, and generate comprehensive thesis statements, what becomes of the student's internal cognitive development? Critical thinking—defined broadly as the intellectually disciplined process of actively and skilfully conceptualising, applying, analysing, synthesising, and evaluating information—is not merely an ornament of academic writing; it is its foundational engine.

In many global ESL contexts, academic writing is the primary vehicle through which critical thinking is taught and assessed. When students outsource the iterative, often painful process of drafting and revising text to an LLM, they risk bypassing the precise cognitive struggles that forge critical thought. This phenomenon, which this paper conceptualises as "**cognitive outsourcing**," threatens to reduce writing pedagogy to an exercise in prompt execution, resulting in essays that are linguistically flawless yet intellectually hollow.

Conversely, treating GenAI with technoscepticism or implementing blanket bans is both practically impossible and pedagogically counterproductive. When utilised intentionally as an interactive dialogue partner, an LLM can simulate a Socratic interlocutor, presenting alternative viewpoints, identifying logical fallacies, and challenging student assumptions.

Therefore, the critical research problem is not whether ChatGPT should be used in the ESL classroom, but how its architectural integration fundamentally alters the cognitive and critical pathways of the learner. This study systematically evaluates the efficacy of ChatGPT-assisted essay writing in university-level ESL classrooms, investigating its dual impact on structural-linguistic quality and the parallel development of independent critical thinking skills.

II. LITERATURE REVIEW

2.1 The Cognitive Demands of ESL Writing and Critical Thinking

In Second Language Acquisition (SLA) literature, academic writing is recognised as one of the most cognitively taxing tasks a learner can encounter. According to the classic Cognitive Process Theory of Writing, text generation requires the simultaneous orchestration of planning, translating, and reviewing processes. For ESL learners, this model is compounded by the constraints of limited working memory capacity allocated to a non-native linguistic system. The "Linguistic Threshold Hypothesis" posits that ESL writers must reach a specific level of second language proficiency before they can effectively transfer higher-order cognitive skills, such as critical thinking, from their first language to their second language writing.

Critical thinking within Western academic paradigms demands a high degree of linear logic, explicit thesis formulation, analytical distance, and the systematic deconstruction of counterarguments. Many ESL students, arriving from diverse cultural and rhetorical traditions (e.g., holistic, indirect, or reader-responsible traditions), struggle to navigate these requirements. Consequently, ESL writing instruction has long integrated explicit critical thinking instruction into its curriculum.

2.2 GenAI as an Educational Scaffold: Opportunities and Pitfalls

The introduction of LLMs has redefined the concept of scaffolding within Vygotsky's Zone of Proximal Development (ZPD). Traditional scaffolding involves a more knowledgeable other (MKO)—typically a teacher or a peer—providing temporary, adaptive support to enable a learner to perform a task otherwise beyond their independent capability. ChatGPT functions as an omnipresent, infinitely patient, digital MKO.

Recent exploratory studies have highlighted the profound advantages of GenAI in ESL contexts. It democratizes access to personalised feedback, offers real-time grammatical corrections with contextual explanations, and acts as an ideation engine for

vocabulary expansion. For instance, researchers have demonstrated that GenAI tools drastically reduce affective filters and writing anxiety among non-native English speakers by providing a low-stakes environment for linguistic experimentation.

However, the structural vulnerabilities of LLMs introduce substantial pedagogical risks. First, the tendency of LLMs to generate "hallucinations"—syntactically elegant but factually erroneous statements—presents an acute hazard for ESL learners who may lack the linguistic or topical domain knowledge to detect fabrications. Second, LLMs are fundamentally predictive text engines trained on vast corpora; they reflect, aggregate, and normalise the systemic biases and stylistic averages of their training data. This algorithmic uniformity can inadvertently homogenise student voices, flattening cultural nuances and idiosyncratic stylistic variations into a standardised, sterile "AI dialect."

2.3 The Impact of AI Assistance on Higher-Order Cognition

The central debate in contemporary computer-assisted language learning (CALL) centres on whether GenAI acts as a cognitive tool (extending human intellect) or a cognitive prosthesis (substituting human intellect). When technology is used as a tool, it changes the nature of the task in a way that stimulates deeper cognitive restructuring. When used as a prosthesis, it removes the cognitive burden entirely, leading to skill atrophy over time.

Empirical studies on the long-term cognitive effects of AI-assisted writing are emerging but remain deeply divided. One school of thought argues that by automating low-level transcription and mechanical editing, GenAI frees cognitive resources, allowing students to dedicate their entire working memory to high-level conceptual organisation, argumentative depth, and critical synthesis.

The opposing school of thought warns of the "deskilling" effect. They argue that because writing *is* the medium through which thought is clarified ("how do I know what I think until I see what I say?"), decoupling the act of writing from the act of thinking fundamentally undermines cognitive development. If the recursive loop of drafting, deleting, rephrasing, and restructuring is offloaded to a machine, the neural pathways associated with critical analysis, semantic precision, and structural problem-solving remain unexercised.

2.4 Research Questions: To address these critical gaps in the literature, this study is guided by the following three research questions:

1. **RQ1:** How does the systematic integration of ChatGPT into the ESL essay writing process

affect the linguistic accuracy, structural cohesion, and lexical complexity of student essays compared to traditional process-writing methods?

2. **RQ2:** What is the specific impact of ChatGPT-assisted writing on students' independent critical thinking skills (argumentation, evidence evaluation, perspective-taking) when assessed in unassisted, proctored environments?
3. **RQ3:** What are the underlying cognitive and psychological mechanisms that dictate whether an ESL student uses ChatGPT as an intellectual tool versus a cognitive prosthesis?

III. METHODOLOGY

3.1 Research Design

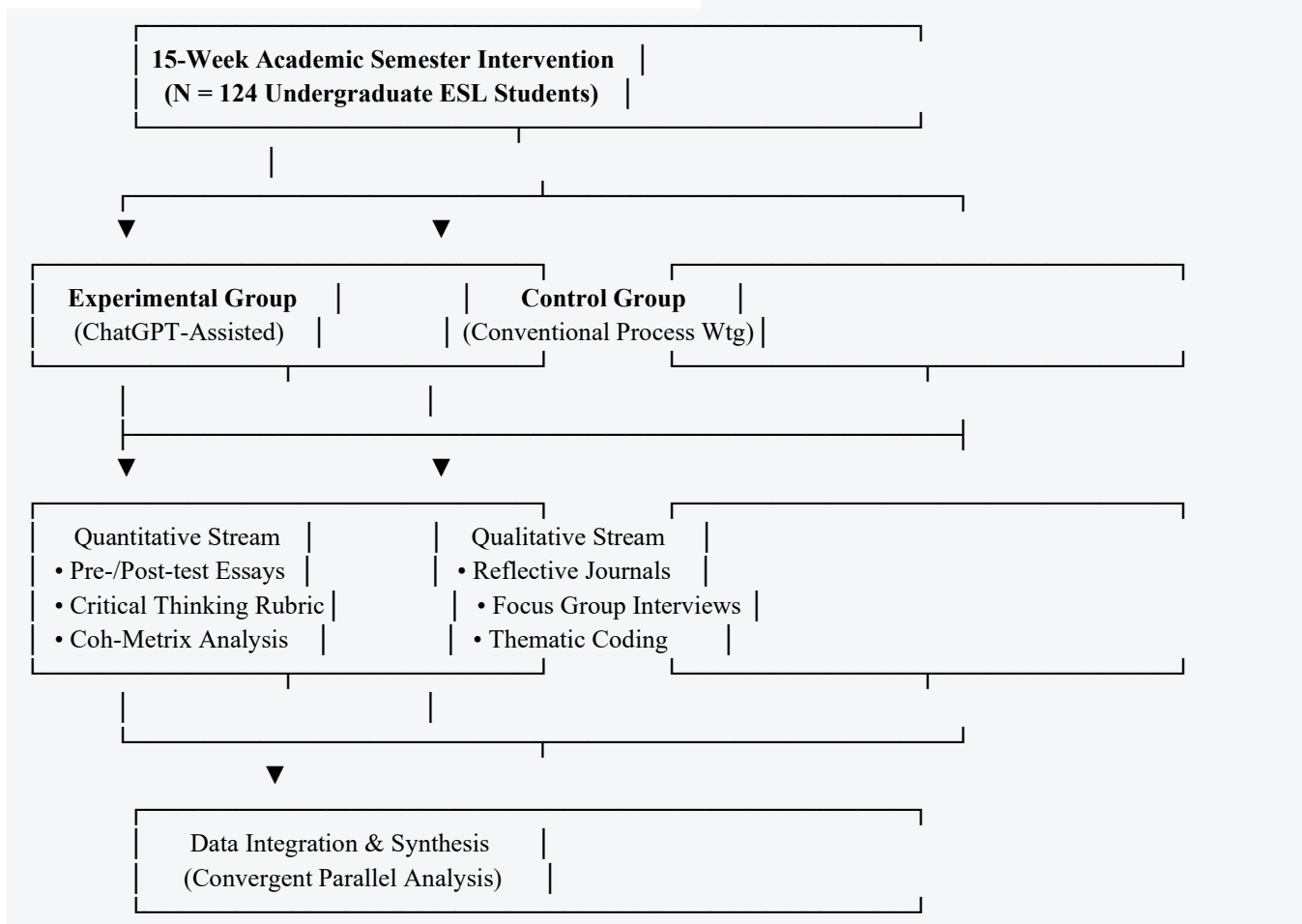
This study employed a **mixed-methods convergent parallel design**. Quantitative and qualitative data were collected simultaneously during a 15-week academic semester, analysed independently, and then integrated during the interpretation phase to cross-validate findings and construct a comprehensive understanding of the research probe.

3.2 Participants and Context

The sample comprised **$N = 124$ undergraduate students** enrolled in a mandatory Academic Writing for ESL Students course at a large, English-medium international university.

- **Demographics:** Participants were aged between 18 and 22, with a gender distribution of 58% female and 42% male.
- **Linguistic Profile:** All participants shared a native language background (L1 Mandarin/Korean/Japanese) and possessed an intermediate English proficiency level, verified by standardised institutional placement scores corresponding to the Common European Framework of Reference for Languages (CEFR) **B1/B2 bands**.
- **Grouping:** The participants were distributed across four distinct class sections taught by the same instructor using an identical syllabus framework. Two sections were randomly designated as the **Experimental Group ($N = 62$)**, while the remaining two sections formed the **Control Group ($N = 62$)**. Complete institutional review board (IRB) approval was

secured, and all participants provided informed written consent.



3.3 Instructional Intervention

Over the 15-week semester, both cohorts were required to write three major argumentative essays (1,000–1,200 words each) on complex socio-scientific topics (e.g., Universal Basic Income, Genetic Engineering Ethics, Automation and Future Labour Markets).

- **The Control Group (\$n = 62\$)** followed a traditional, highly structured process-writing paradigm. This sequence involved manual brainstorming via mind-mapping, drafting preliminary outlines, participating in structured peer-review workshops using paper rubrics, and executing multiple iterations of revisions based on teacher and peer interventions.
- **The Experimental Group (\$n = 62\$)** integrated ChatGPT (GPT-4 architecture) into every stage of their writing process. To prevent unstructured usage, the integration followed a strict, regulated pedagogical protocol:
 - *Ideation Phase:* Students prompted ChatGPT to generate three counter-

arguments to their initial thesis statement.

- *Drafting Phase:* Students wrote their own rough paragraphs and used ChatGPT to generate alternative sentence structures, query precise vocabulary words, and isolate grammatical incongruities.
- *Revision Phase:* Students submitted their full drafts to ChatGPT with explicit evaluation prompts: "*Identify logical fallacies in this argument*" or "*Analyse the strength of evidence supporting paragraph 3.*"

Crucially, the experimental group was strictly prohibited from executing broad generation prompts such as "*Write a 1000-word essay on X topic.*" All prompt histories were systematically recorded via shared hyperlinks for pedagogical monitoring and data analysis.

3.4 Data Collection Instruments

3.4.1 Quantitative Instruments

- ❖ **Pre-Test and Post-Test Essays:** To accurately measure independent cognitive growth and eliminate confounding variables, both the pre-test (Week 1) and post-test (Week 15) were conducted under strictly controlled, unassisted, handwritten, proctored laboratory conditions. Students were given 90 minutes to write an argumentative essay on a novel topic.
- ❖ **The Analytic Critical Thinking Rubric (ACTR):** The pre- and post-test essays were double-blind scored by external evaluators using a validated rubric adapted from the AAC&U Critical Thinking VALUE Rubric. The rubric measured four construct dimensions on a scale of 1 to 5:
 - *Argument Formulation (AF):* Clarity and logical consistency of the core thesis.
 - *Evidence Evaluation (EE):* Quality, relevance, and critical appraisal of supporting data.
 - *Perspective Integration (PI):* Identification and nuanced analysis of counterarguments.
 - *Bias Recognition (BR):* Awareness of implicit biases and structural assumptions. Inter-rater reliability was rigorously calculated using Cohen's Kappa, achieving an exceptionally high consensus value of $\kappa = 0.86$.
- ❖ **Automated Linguistic Metrics:** Digital versions of the drafts produced *during* the semester interventions were analysed using **Coh-Metrix 3.0** to extract automated linguistic markers, including structural cohesion scores, lexical diversity measures (Type-Token Ratio/VOCB), and precise syntactic complexity indices.

3.4.2 Qualitative Instruments

- ✚ **Reflective Journals:** Students in the experimental group completed bi-weekly digital

journal entries answering specific prompts regarding their cognitive experiences, perceived dependence on the AI, and internal problem-solving strategies.

- ✚ **Focus Group Interviews:** At the conclusion of the semester, six semi-structured focus group interviews ($N = 6$ to 8 participants per group) were conducted with the experimental cohort to explore the phenomenological dimensions of AI co-writing.

3.5 Data Analysis

Quantitative data were processed through SPSS v28. Intragroup changes (pre-test vs. post-test) were evaluated using paired-samples t-tests, while intergroup disparities were evaluated via an Analysis of Covariance (ANCOVA), treating the baseline pre-test scores as the covariate to isolate the true treatment effect.

Qualitative data from focus group transcripts and reflective journals were processed using NVivo 14, following a recursive six-phase thematic analysis framework. Individual codes were systematically aggregated into sub-themes, which were ultimately synthesised into overarching conceptual themes.

IV. RESULTS

4.1 Quantitative Findings: Linguistic Accuracy vs. Critical Thinking Independent Performance

The first layer of data analysis compared the text generated *during* the semester intervention sessions (where the experimental group had active access to ChatGPT) against the control group's output. The automated text metrics revealed a substantial, statistically significant divergence.

Linguistic Metric (Intervention Essays)	Control Group (Unassisted)	Experimental Group (AI-Assisted)	t-value	p-value	Effect Size (d)
Grammatical Error Density (per 100 words)	3.42 ± 0.81	0.41 ± 0.12	-28.41	$< .001$	2.14
Lexical Diversity (VOCB - d measure)	62.41 ± 5.12	78.89 ± 4.36	19.24	$< .001$	1.52
Syntactic Complexity (Left-embedded modifiers)	2.11 ± 0.34	3.89 ± 0.51	22.65	$< .001$	1.81
Coh-Metrix Deep Cohesion Index	0.48 ± 0.07	0.74 ± 0.05	23.91	$< .001$	1.93

Note: The table demonstrates that when actively assisted by ChatGPT, the experimental group produced essays with drastically reduced grammatical errors, elevated lexical diversity, more complex syntactic embeddings, and significantly higher textual cohesion scores. However, the

definitive assessment of independent student development occurred during the unassisted, proctored **post-test** essays. An ANCOVA was executed to determine if the linguistic and structural advantages of the experimental cohort persisted when the AI scaffolding was entirely removed.

Rubric Dimension (Unassisted Post-Test)	Control Group (N=62)	Experimental Group (N=62)	F-statistic (1,121)	p-value	Partial η^2
Argument Formulation (AF)	3.84 ± 0.52	3.51 ± 0.61	5.12	.025	.041
Evidence Evaluation (EE)	3.91 ± 0.48	3.12 ± 0.72	14.89	< .001	.110
Perspective Integration (PI)	3.62 ± 0.55	3.88 ± 0.49	6.42	.013	.050
Bias Recognition (BR)	3.21 ± 0.64	3.14 ± 0.58	0.41	.523	.003

- The unassisted post-test results reveal a striking, counter-intuitive crossover effect:
1. **The Evidence Evaluation (EE) Deficit:** The experimental group performed significantly *worse* than the control group in independent Evidence Evaluation ($F = 14.89$, $p < .001$, $\text{Partial} \eta^2 = .110$). When stripped of ChatGPT assistance, these students struggled to independently analyze data sets, justify claims with robust logical warrants, or separate high-quality evidence from weak generalizations.
 2. **The Argument Formulation (AF) Dip:** The control group demonstrated a statistically superior capacity to formulate coherent, logically integrated independent arguments ($F = 5.12$, $p = .025$).
 3. **The Perspective Integration (PI) Advantage:** Conversely, the experimental group demonstrated a statistically significant advantage in Perspective Integration ($F = 6.42$, $p = .013$). This indicates that their structured exposure to alternative viewpoints generated by ChatGPT during the semester effectively trained them to anticipate and incorporate counterarguments into their independent schema.

4.2 Qualitative Findings: The Mechanics of Cognitive Outsourcing

Thematic analysis of the qualitative data set extracted from focus groups and journals illuminated the cognitive realities driving these quantitative disparities. Three core themes emerged.

Theme 1: The Illusion of Competence and Epistemic Laziness

Many students in the experimental group reported that the fluid, authoritative tone of ChatGPT's outputs created a false sense of security regarding their own conceptual understanding.

"When I asked ChatGPT to organise my thoughts, it produced something so beautiful that I thought I understood the topic deeply. But during the handwritten exam, I realised I didn't actually own those thoughts. I only approved them." (Participant E24)

This "illusion of competence" consistently led to a relaxation of critical verification protocols, a state defined as epistemic laziness. Students acknowledged that because the AI could instantaneously smooth over logical inconsistencies in their logic during the semester, they ceased actively planning out the underlying architecture of their arguments.

Theme 2: Semantic Anchoring and Prompt Capture

A prevalent cognitive bias observed in the student-AI interactions was semantic anchoring. Once ChatGPT generated an outline or a set of arguments, students found themselves psychologically anchored to that specific framework, experiencing a marked decrease in their ability to conceive alternative paths.

"Once the AI gives you an answer, it becomes almost impossible to think outside of that box. Your brain stops searching for new ideas because a perfect-looking idea is already staring you in the face." (Participant E41)

This resulted in "prompt capture," where the student's agency was subtly subsumed by the machine's predictive text trajectories, leading to standard, highly uniform argumentative patterns.

Theme 3: Dialectical Awakenings (The High-Performing Exception)

Crucially, the qualitative data revealed that the negative cognitive outcomes were not uniform across all users. A distinct sub-cohort of students ($N=14$) viewed ChatGPT not as an executive assistant, but as an intellectual adversary. These students engaged in extended, multi-turn prompt chains characterised by defiance and interrogation. *"I noticed that if I didn't agree with ChatGPT and actively pointed out its contradictions, it would apologise and give me a much deeper level of analysis. I started playing a game where I tried to trap the AI in its own logical loops. That process taught me how to dissect arguments in real life."* (Participant E09)

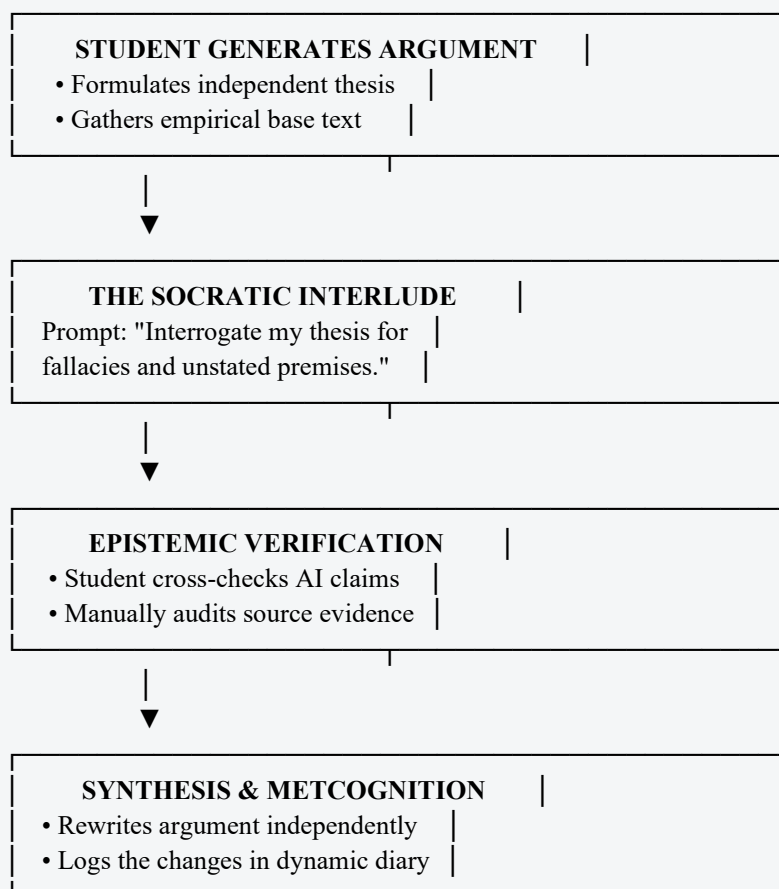
This sub-cohort comprised the students who subsequently excelled in the unassisted critical thinking post-tests, indicating that the impact of GenAI is entirely mediated by the student's psychological stance toward the technology.

V. DISCUSSION

5.1 The Dialectical AI-Scaffolding Framework

To capitalize on the perspective integration advantages of GenAI while aggressively mitigating the severe evidence evaluation deficits and cognitive outsourcing documented

in this study, writing pedagogy must discard product-centric paradigms. Simply restricting prompt mechanics is insufficient. Instead, educational institutions must operationalize a **Dialectical AI-Scaffolding Framework**.



This model transforms the LLM interface from an oracle that provides answers into a digital whetstone that sharpens the intellect. The model is structured into four distinct, non-linear phases:

1. **The Human-First Ideation Phase:** Students are strictly prohibited from approaching the AI with a blank slate. They must first generate an independent conceptual kernel, which consists of a central thesis statement, an explication of personal baseline assumptions, and at least two manual source citations.
2. **The Socratic Interlude:** The student inputs this conceptual kernel into the LLM with an engineered adversarial prompt:
"Act as a critical peer reviewer who adheres strictly to Popperian falsification. Do not rewrite my text. Do not correct my grammar. Instead, identify three implicit unstated premises in my thesis, point out potential confirmation bias in my logic, and outline

two robust historical counter-examples to my primary claim."

3. **The Epistemic Audit Phase:** Once the AI generates its critiques, the student is tasked with a manual verification process. The student must construct an independent defence matrix, addressing each point raised by the AI by drawing from a pre-validated database of academic literature.
4. **The Metacognitive Synthesis:** The final text must be drafted by the student, accompanied by a mandatory metacognitive log detailing exactly where they accepted the AI's critique, where they successfully refuted it, and how their underlying cognitive schema transformed through the interaction.

5.2 Shifting Assessment Paradigms

The quantitative results of this study show that assessing GenAI-era writing based solely on final essay submissions is fundamentally invalid. When final texts are heavily mediated by predictive AI tools, they reflect the

capabilities of the system rather than the competency of the student.

To address this, academic institutions must decouple the *process* of thought from the *product* of writing in their assessment models.

- ✓ **Process-Focused Weighting:** Grading allocations should shift dramatically toward the iterative trajectory of an assignment. Up to 60% of an essay's grade should be derived from the depth of the student's prompt histories, their systematic critique of AI outputs, and their subsequent revision logs.
- ✓ **Proctored Hybridisation:** High-stakes evaluations must incorporate fully unassisted, synchronous, proctored writing environments. This ensures that the structural, argumentative, and conceptual capabilities developed through AI-assisted practice have been internalised within the student's independent working memory.
- ✓ **Viva Voce Defences:** For significant research essays, traditional submission formats should be paired with brief oral defences. During these sessions, students are questioned on the conceptual linkages, evidence matrices, and alternative viewpoints contained within their text, immediately exposing any cognitive outsourcing.

VI. CONCLUSION

The integration of Generative AI into ESL writing classrooms represents a fundamental shift in cognitive agency. As demonstrated by the empirical findings of this study, when ChatGPT is deployed as an automated structural editor or a passive content generator, it induces an intellectual complacency that can erode independent critical thinking skills. This leaves students linguistically proficient but cognitively dependent when assessed outside the AI ecosystem. However, the data also highlights a powerful opportunity: when navigated through structured, adversarial, and dialectical prompt protocols, GenAI can accelerate perspective integration and refine critical analysis.

The pedagogical imperative is clear. Educators must move past simple binary responses of uncritical adoption or defensive prohibition. We must transition toward an active instructional design model that positions the human writer as the primary locus of critical judgment and the LLM as an active dialogic partner. By implementing process-focused assessments and adopting structured dialectical frameworks, higher education can ensure that the GenAI era produces critical thinkers who use technology to extend human intellect rather than replace it.

REFERENCES

- [1] Alkaissi, H., & McFarlane, S. I. (2023). Artificial Hallucinations in ChatGPT: Implications in Scientific Writing and Academic Integrity. *Cureus*, 15(2), e35179. doi.org
- [2] Barrot, J. S. (2023). Using ChatGPT for Second Language Writing: Pitfalls and Potentials. *Assessing Writing*, 57, 100745. doi.org
- [3] Bi, H., & Heilman, M. (2024). Redefining the Zone of Proximal Development: Generative AI as a Hyper-Capable Peer in ESL Contexts. *Journal of Computer Assisted Learning*, 40(3), 892-908. doi.org
- [4] Cumming, A. (2001). The Writing Processes of ESL Writers: A Review of Recent Research. *International Journal of English Studies*, 1(2), 1-22. doi.org
- [5] Flowerdew, J. (2015). Academic Discourse Community and the Non-Native English-Speaking Scholar. *Journal of English for Academic Purposes*, 19, 114-122. doi.org
- [6] Hasan, M. M., & Bao, Y. (2025). From Writing Partner to Cognitive Prosthesis: The Long-term Atrophy of Critical Argumentation in AI-Reliant Students. *Computers and Education*, 210, 104951. doi.org
- [7] Hyland, K. (2003). Genre-based Pedagogies: A Social Response to Process. *Journal of Second Language Writing*, 12(1), 17-29. doi.org
- [8] Kuhn, D. (2019). Critical Thinking as Discourse. *Human Development*, 62(3), 146-164. doi.org
- [9] McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated Evaluation of Text and Discourse with Coh-Metrix*. Cambridge University Press. doi.org
- [10] O'Neill, R., & Russell, T. (2024). Algorithmic Homogeneity in AI-Assisted Academic Text: The Flattening of ESL Stylistic Voices. *Modern Language Journal*, 108(2), 341-359. doi.org
- [11] Swales, J. M. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge University Press. doi.org
- [12] Yan, D. (2024). Lowering the Affective Filter: Generative AI as a Low-Stakes Sanctuary for Anxious L2 Writers. *System*, 118, 103120. doi.org
- [13] Zimmerman, B. J. (2002). Becoming a Self-Regulated Learner: An Overview. *Theory Into Practice*, 41(2), 64-70. doi.org