Research Article

IMPLICATION OF GEN-AI-INTEGRATED ELT INSTRUCTION ON STUDENTS' CRITICAL THINKING IN ACADEMIC WRITING

*Bethsua I. Bazar, Mark Louie C. Huraño and Fe R. Janiola

Graduate Studies, Holy Name University, Tagbilaran City, Bohol, Philippines

Abstract

This study investigated the implication of Generative AI (GenAI)-Integrated ELT Instruction on the critical thinking in academic writing of Grade 12 TVL and ABM students at Hinlayagan National High School. Using an embedded mixed-methods research design, the researcher compared a traditional English Language Teaching (ELT) method with a GenAI-integrated ELT instruction. Quantitative analysis through T-test results revealed a significant performance gap: while the traditional group demonstrated substantial gains, the experimental group showed a measurable decline in critical thinking in academic writing. Qualitative findings identified a "productivity paradox," where students utilized GenAI to bypass the "time tax" of drafting, resulting in skill atrophy. This over-reliance led to "cognitive debt," where the ease of the technology caused a loss of original reflection and independent critical thinking in answering essay writing tasks.The study concludes that without a structured pedagogical framework in integrating GenAI may lead to cognitive stagnation, GenAI tends to substitute for rather than supplement human logic. To mitigate these risks, the research proposes the Critical-Scaffolding (CS) Framework, which redefines the teacher's role as a critical evaluation coach. This framework advocates for "GenAI-required" tasks that align with DepEd standards to ensure mandatory cognitive engagement and the protection of authentic academic authorship.

Keywords: Generative AI, ELT Instruction, Critical Thinking, Academic Writing, Critical-Scaffolding Framework, Productivity Paradox.

INTRODUCTION

The emergence of generative artificial intelligence (Gen AI) marks a significant turning point in technological progress across many fields. This technological leap has deeply influenced various sectors, including education, and specifically impacts how languages are taught and learned. In English language teaching (ELT) classrooms, AI tools are continually evolving in complexity and functionality. Therefore, their potential impact on students' learning experiences and academic performance warrants careful examination. A key concern arising from this integration is precisely how Gen AI (especially tools like conversational LLMs) is shaping students' critical thinking, particularly within the context of academic writing. Critical thinking in academic writing is essential for students in today's rapidly changing educational landscape. Students must be equipped to analyze information, evaluate sources effectively, and construct sophisticated logical arguments. However, to effectively convey these critical-thinking skills, students must first possess a strong command of the English language. This linguistic proficiency forms the necessary foundation for articulating complex ideas, organizing thoughts coherently, and utilizing an appropriate academic tone and vocabulary. Fostering English language mastery is thus a fundamental and crucial step in empowering students to engage deeply in critical thinking and produce compelling academic work. On one hand, the integration of generative AI into ELT is increasingly vital for basic education institutions as they respond to the growing influence of technology. As schools worldwide adopt AI tools to enhance English instruction (Crompton, 2024; Edmett, 2023), it becomes crucial to fully understand how Gen AI can affect students' critical thinking in academic writing. Incorporating Gen AI can either enhance or potentially hinder the overall development of critical thinking skills. It aids those facing language barriers by improving grammar, phrasing, and vocabulary allowing students to concentrate more on content. Conversely, over-reliance on AI may inhibit the practice and refinement of students' independent language composition skills, thus hindering academic writing development (Lee *et al.,* 2025; Liu *et al.,* 2025). The increasing use of AI tools by students globally reflects a significant and ongoing shift toward embracing advanced technology for academic success. Some studies report that students are off-loading higher-order cognitive tasks to AI, raising concerns about diminished critical thinking (Wardle, 2025). This pronounced trend indicates that modern education is progressively incorporating AI to support learning processes and simplify various academic tasks. This broad adoption underscores the need to study its specific implications on critical thinking development. Furthermore, governmental bodies are now actively advocating for the responsible integration of artificial intelligence into school systems. For example, the Department of Education (Philippines) (DepEd) has publicly emphasized the use of AI in schools, although peer-reviewed empirical research on this local context remains scarce. While AI-based tools are gaining traction and aiding academic success globally, there is a concerning local trend of declining critical thinking among Filipino students—largely attributed to the sustained dominance of rote memorization within the existing Philippine educational system (Dela Cruz, 2012; Lugtu, 2019). In the local context, a pre-survey conducted among senior high school students revealed a significant challenge: those who did not use Gen AI demonstrated low levels of critical thinking in academic writing. This data strongly suggests that students have an existing difficulty in independently formulating original ideas

*Corresponding Author: *Bethsua I. Bazar,*
Graduate Studies, Holy Name University, Tagbilaran City, Bohol, Philippines.

and crafting well-supported essay arguments. Therefore, reliance on AI-generated content can potentially further weaken students' ability to develop their own critical thinking and original arguments. Providing effective English language instruction, particularly focused on academic writing, is thus essential for fostering students' critical thinking capacity. However, incorporating generative AI into these academic environments clearly presents various practical and pedagogical challenges and concerns. There is currently a significant gap in structured guidance for educators navigating this new technological landscape (British Council, 2024). Thus, this study fundamentally aims to establish a clear, evidence-based teaching framework. This framework will guide educators and policymakers in creating effective, responsible guidelines for the measured and beneficial integration of generative AI in ELT instruction.

Objectives of the Study

This study aims to determine the effect of Gen-AI-Integrated ELT instruction on the students' critical thinking in academic writing. This will craft a teaching framework that promotes students' critical thinking in academic writing despite the use of Gen AI.

Specifically, this study seeks to find answers to the following questions:

1. What is the student's performance in critical thinking in academic writing, as measured by their scores in the:
   - Pre-Test and Post-Test of the experimental group (Gen AI-Integrated instruction)?
   - Pre-Test and Post-Test of the control group (Traditional Instruction)?
2. Is there a significant difference of the student's performance in the critical thinking in academic writing when comparing the:
   - 2.1 Pre-Test and Post-Test scores within the control group?
   - 2.2 Pre-Test and Post-Test scores within the experimental group?
   - 2.3 Post-test scores between the control group and experimental group?
3. What are the perceived critical thinking challenges faced by students in academic writing when utilizing Generative AI compared to those who solely rely on conventional methods?
4. What are the pedagogical implications of integrating Generative AI on ELT instruction in the classroom?
5. What teaching framework for the responsible and effective integration of Generative AI in ELT instruction can be proposed based on the findings of the study?

METHODOLOGY

The research on the implication of Gen-AI-integrated ELT instruction on students' critical thinking in academic writing employed a mixed methods research design to effectively accomplish its objective. This comprehensive approach allowed for the simultaneous collection and thorough analysis of both qualitative and quantitative data types. Integrating these findings and drawing inferences was achieved through the use of diverse methodologies within a single, cohesive program of inquiry (Tashakkori & Creswell, 2007).

Specifically, the design's purpose, as outlined by Creswell and Clark (2014), was to merge the individual results from both data strands during the interpretation phase. Ultimately, this synergy determined how the datasets converged or diverged, providing a more comprehensive basis for formulating informed policy regarding the use of Gen-AI integration in ELT instruction within the institution. The quantitative phase of this study utilized a quasi-experimental design employing a Pre-Test and Post-Test structure to analyze measurable data generated through experimentation. This specific method was crucial because it allowed the researcher to evaluate the causal impact of the Gen-AI integration intervention. Simultaneously, the researcher also incorporated qualitative inquiry to supplement the experiment with rich, observational data and relevant accounts. These 'soft data' were essential for gaining deeper insights into the process, going beyond the numbers. The qualitative data extracted directly informed and guided the final design of the proposed teaching framework. The core of the study's experimentation involved a setup with both a control group and a distinct experimental group to facilitate a clear comparison of outcomes. Following the implementation of the intervention and the collection of quantitative data, the researcher conducted in-depth interviews with the participants. These interviews were designed to generate detailed qualitative information concerning students' performances with and without generative AI utilization in their academic writing. Furthermore, the interviews sought information about the practices and perspectives of teachers regarding their use of Gen-AI in ELT instruction. This overall methodology aligned with an embedded experimental design (Creswell & Clark, 2014), which characteristically used one method to enhance or support the primary data (the quantitative experiment).

RESULTS

Student's Performance in Critical Thinking in Academic Writing in Pre-Test and Post-Test across Groups

Student's Performance Using Gen-AI-Integrated Instruction Experimental Group: Gen-AI-Integrated ELT Instruction

Table 1. Pre-test and Post Test of the Experimental Group Using Gen-AI-Integrated ELT Instruction

Range	Description	Pre-Test		Post-test	
		Frequency (f)	Percentage (%)	Frequency (f)	Percentage (%)
39-50	High	-	-	1	3%
26-38	Moderate	4	13%	1	3%
13-25	Low	5	17%	2	7%
0-12	Very Low	21	70%	26	87%
Total		30	100%	30	100%
Average Score		9.4		5	
SD		4.18		4.18	
Description		Very Low		Very Low	

Table 1 presents the pre-test and post-test performance of students exposed to Gen-AI-Integrated ELT Instruction. In the pre-test, the majority of students 21 or 70%scored within the Very Low performance range of (0–12). Seventeen percent of students were classified as Low and there are 13% at Moderate level. The group obtained a mean score of 9.4 during the pre-test.

The SD remained consistent at 4.18 across both assessments, which is key to interpret the overall group outcome. The quantitative result has a substantial and significant decline in critical thinking performance. The experimental group's mean score decreased from 9.4 in the pre-test down to 5.0 in the post-test phase.

This considerable drop provides clear evidence of a significant negative impact or, at minimum, a definitive failure of the instruction to teach the targeted critical thinking in academic writing. Analysis confirms that the students' performance remained fixed at the most deficient level of accomplishment throughout the study. Since both the pre-test mean (9.4) and the post-test mean (5.0) fall within the Very Low range (0–12), the Gen AI-integrated instruction failed to raise performance above the minimum categorical level. Contrary to all initial expectations, the Gen-AI-integrated instruction did not enhance critical thinking in academic writing; instead, it produced the opposite effect on the experimental group. This negative outcome suggests students may have relied on the AI or were discouraged from active engagement, leading to decreased critical thinking in academic writing. This decline is notable when contrasted with the control group, which, receiving traditional instruction, demonstrated natural improvement during the same period. Moreover, another study the seminal "Your Brain on ChatGPT" study by Kos'myna et al. (2025) provides the essential foundation for Gen AI's cognitive implications, demonstrating through EEG data that sustained Large Language Model use leads to weaker neural connectivity during essay composition, thereby confirming the neurological outsourcing of critical thought and establishing a vital link to the risk of critical thinking atrophy.

Student's Performance Writing Using Traditional Method
Control Group: Traditional Method

Table 2 shows the pre-test and post-test performance of students who received traditional instruction. During the pre-test, the majority of students (25 out of 30, 83%) scored in the Very Low range. Three students (10%) were classified as Low, and two students (7%) were classified as Moderate. The group obtained a mean score of 11.0 (SD = 4.90), corresponding to a Very Low level. In the post-test, the distribution of scores shifted across higher performance levels. Six students (20%) attained High scores, thirteen (43%) were classified as Moderate, and nine (30%) fell under the Low category. Only two students (6%) remained in the Very Low range. The post-test mean score increased to 26.0 with a standard deviation of 7.92, corresponding to a Moderate level of performance.

The increase in standard deviation indicates greater variability in post-test scores compared to pre-test scores.

Moreover, as reflected in table 2, the analysis of the control group, which received instruction through the traditional method, reveals a substantial and positive shift in student performance regarding critical thinking. The group's average score rose from an initial Pre-Test mean of 11 to an impressive Post-Testmean of 26, representing a 15-point gain that elevated the achievement level from Very Low to Moderate. This strong quantitative improvement confirms the inherent efficacy of conventional teaching strategies in enhancing critical thinking and establishes an essential baseline of expected positive improvement. This transformation of students' performance in critical thinking in academic writing is further illustrated by the decisive shift in frequency across the established achievement ranges. At baseline, the majority of students (83%, f=25) were clustered in the Very Low category, but this group collapsed to only 6% (f=2) following traditional instruction. Also, the number of students achieving a Moderate level (26-38) surged from 7% to a strong majority of 43%, with an impressive 20% of the group also reaching the High achievement level (39-50).

This positive redistribution of scores highlights the success of the traditional method in fostering foundational and advanced critical thinking in writing across the student body. A key finding in the control group's data is the distinct increase observed in the Standard Deviation (SD) between the two assessments. The SD rose from 4.90 in the Pre-Test to 7.92 in the Post-Test, which shows a greater variability in the distribution of student scores. This observed increase in variability indicates that, even with an overall group improvement, the size of individual student gains differed significantly across the participants. It is evident that the traditional English Language Teaching (ELT) instruction elevated the group's average performance in academic writing, yet this success created a wider spread of achievement by highlighting differential learning paces among participants. Overall, these metrics affirm the effectiveness of the established teaching method while emphasizing the varied individual responses and learning rates observed throughout the study period.

This finding aligns with the study by Tsui, L. 2002. Tsui's findings provided strong empirical support that traditional, active learning methodologies implemented within writing are the most reliable way to cultivate these necessary cognitive abilities. It provides a clear explanation of *why* the traditional writing method is effective in promoting these higher-order skills. Requiring students to construct independently a detailed academic essay fundamentally forces them to engage in the necessary cognitive processes which are the hallmarks of critical thought.

Table 2. Pre-Test and Post-test of the Control Group Using Traditional Method

Range	Description	Pre-test		Post-test	
		Frequency (f)	Percentage (%)	Frequency (f)	Percentage (%)
39-50	High	-	-	6	20%
26-38	Moderate	2	7%	13	43%
13-25	Low	3	10%	9	30%
0-12	Very Low	25	83%	2	6%
Total		30	100%	30	100%
Average Score		11		26	
Standard Deviation		4.90		7.92	
Description		Very Low		Moderate	

Table 3. Pre-test Performance of the Control and Experimental Groups

Range	Description	Control Group		Experimental Group	
		Frequency (f)	Percentage (%)	Frequency (f)	Percentage (%)
39-50	High	-	-	1	3%
26-38	Moderate	2	7%	1	3%
13-25	Low	3	10%	2	7%
0-12	Very Low	25	83%	26	87%
Total		30	100%	30	100%
Average Score		11		9.4	
SD		4.90		4.18	
Description		Very Low		Very Low	

The pre-test results as shown in Table 3 indicates that both the control and experimental groups started with very low performance in critical thinking in academic writing. In the control group, the majority of students 83% (25 out of 30, achieved scores in the Very Low range (0-12), with only a few reaching Low (10%) or Moderate (7%) levels. Their average score was 11 (SD = 4.90), confirming the group's overall Very Low performance level. The experimental group showed a similar result before the intervention began. Seventy percent students (21 out of 30) of these students also fell in the Very Low range on pre-test, with small portions achieving Moderate (17%) or High (17%) performance, averaging 9.4 (SD = 4.18), also classified as Very Low. These parallel results establish a comparable baseline for both groups before the introduction of the GenAI integrated instruction. Both the control and experimental groups began with a Very Low average performance, confirming that they both lacked the necessary critical thinking for academic writing and establishing no initial advantage for the experimental group prior to the Gen AI intervention.

Table 4. Post-test students' Performance of the Control and Experimental Group

Range	Description	Control Group		Experimental Group	
		Frequency (f)	Percentage (%)	Frequency (f)	Percentage (%)
39-50	High	6	20%	1	3%
26-38	Moderate	13	43%	1	3%
13-25	Low	9	30%	2	7%
0-12	Very Low	2	6%	26	87%
Total		30	100%	30	100%
Average Score		26		5	
Standard Deviation		7.92		4.18	
Description		Moderate		Very Low	

Table 4 shows the post-test results between the control and experimental groups, indicating a substantial difference in critical thinking in academic writing between the control and experimental groups. The control group shows a clear and positive improvement, with 6 students (20%) achieving High scores, 13 (43%) Moderate, and 9 (30%) Low, while only 2 students (6%) remain in the Very Low range, resulting in an average score of 26 (SD = 7.92), moving their overall performance into the Moderate classification. In contrast, the experimental group's performance declined sharply following the intervention, with 26 students (87%) still in the Very Low range, only 2 students (7%) at Low, and 1 student (3%) each in Moderate and High ranges, yielding anaverage score of 5 (SD = 4.18), indicating continued Very Low performance. Contrary to expectations, the data indicates that the Gen-AI-integrated instruction failed to enhance critical thinking in academic writing, likely due to students' over-reliance on the AI or discouraged independent engagement, which led to a decrease in their overall performance. The findings of the study demonstrate that the traditional method improved the critical thinking in academic writing of the control group, while the experimental group's performance either stagnated or deteriorated, underscoring the potential negative implications of Gen AI integration on essential academic competencies. The pre-test and post-test comparison of the control group shows a significant improvement in critical thinking in academic writing. The control group's mean score increased from 11 ± 4.90 to 26 ± 7.92, and the t-test revealed a significant difference (t = -9.32, p = 0.003). This indicates that traditional instruction was effective in enhancing students' performance in critical thinking in academic writing, allowing most students to progress from very low to moderate performance. The negative t-value reflects this positive change over time. The experimental group exposed to Gen-AI-integrated instruction demonstrated a significant decline in performance. Their mean score decreased from 11 ± 4.90 to 5 ± 4.18, with a t-value of 2.20 and p = 0.03, indicating a significant change. However, this change was negative, suggesting that AI integration did not support students' development of critical thinking in academic writing. Instead, it may have hindered active engagement, resulting in lower post-test scores. When comparing pre-test scores between the control and experimental groups, there was no significant difference (t = 1.35, p = 0.18), confirming that both groups started at a comparable baseline. This ensures that the observed differences in post-test outcomes are attributable to the type of instruction rather than initial disparities in ability.

The post-test comparison between the groups shows a significant difference (t = 10.73, p = 0.000), with the control group outperforming the experimental group by a wide margin. This emphasizes that traditional instruction was far more effective in fostering critical thinking in academic writing than the Gen-AI-Integrated. Overall, the research findings suggest thatan excessive dependency on Generative AI within English Language Teaching can suppress a student's capacity for independent critical thought. Furthermore, this overreliance leads to a measurable decline in overall students' performance in critical thinking in academic writing highlighting the need for careful scaffolding, monitoring, and strategic integration of GenAI in educational settings. To understand the unexpected quantitative decline, the researcher started the qualitative phase by conducting in-depth interviews. These interviews included students from the GenAI experimental group and the English subject teachers involved in the instruction. Following the careful analysis of these comprehensive interview transcripts, the researcher synthesized the key emergent data into two major interconnected thematic categories for discussion that explain the perceived critical thinking challenges.

Perceived critical thinking challenges in academic writing when utilizing generative ai compared to conventional methods

The overall perceived critical thinking challenges are a composite of the benefits offered by AI and the burdens associated with conventional writing. The first theme focuses on the overwhelming utility of Gen AI as the preferred method for bypassing these existing academic barriers.

GenAI as a mechanism for efficiency and linguistic enhancement

This theme focuses on the benefits of Gen AI, which outweigh the effort required by traditional school work, highlighting the overwhelming practical utility of Generative AI (Gen AI). This substantial utility and the capacity for immediate results are what makes the Gen AI method the preferred choice among students for task completion. However, students perceive the challenge of conventional methods as being time, effort, and linguistic deficit. Gen AI resolves these barriers, making the "hard work" of critical thinking obsolete. Gen AI solves the major problem students face: the drain on their time and effort when completing school work. Students chooseAI because it lets them bypass the traditional academic hassle ofconsuming too much time and feeling exhausted.

Table 8. Difference in the Post-test of the Experimental and Control Group

Variables	Mean$\pm$*SD*	Computed T-value	P-value	Decision on Ho	Interpretation
Pre- Post Test Control Group	11 $\pm$ 4.90 9.4 $\pm$ 4.18	-9.32	0.003	Reject Ho	Significant

Table 7. Difference in the Pre-test of the Experimental and Control Group

Variables	Mean$\pm$*SD*	Computed T-value	P-value	Decision on Ho	Interpretation
Pre- Post Test Experimental Group	11 $\pm$ 4.90 5.0 $\pm$ 4.18	2.20	0.03	Reject Ho	Significant

Table 6. Difference in the Pre-test and Post-test of the Experimental Group

Variables	Mean$\pm$*SD*	Computed T-value	P-value	Decision on Ho	Interpretation
Pre-test Experimental & Control Group	11 $\pm$ 4.90	1.35	0.18	Accept Ho	Not Significant

Table 5. Difference in the Pre-test and Post-test of the Control Group

Variables	Mean$\pm$*SD*	Computed T-value	P-value	Decision on Ho	Interpretation
Post-test Experimental & Control Group	5 $\pm$ 4.18 26 $\pm$ 7.9	10.73	0.000	Reject Ho	Significant

Discussant 1 explained that the biggest risk of not using AI is that it is "time consuming," showing they see AI as the essential fix for achieving "speed and effortless work." In addition to speed, Gen AI gives students high-quality work right away, helping them with their writing structure and new vocabulary. Students turn to the tool to guarantee a successful academic product, knowing it will meet high standards and avoid grammar errors. Discussant 3 used the AI to "make a perfect performance in making essay," showing how the tool overcomes the perceived challenge of having insufficient writing skills or vocabulary.

In essence, Discussant 1, 2 & 3 stated that the perceived challenge of conventional writing is that it is "time consuming" and leaves their brain "exhaust or drained". Because traditional work is so taxing, students use AI to transfer the effort away from themselves. This "offloading simplifies the entire problem-solving process required to create complex academic responses, making AI the clear choice. Crucially, the most important benefit students mention is the tool's ability to help them finish the tasks. Discussant 1 & 3 reported that they "can answer question fast" and "can make right away my performance tasks" using AI. This behavior highlights a clear choice: students prioritize immediate speed and completion over the deliberate, slower process of independent thinking required for genuine learning. Therefore, students cited speed and effortlessness as the motivators for Gen AI use. Discussant 1 &3 responses such as "we can answer question fast" and the ability to "make right away my performance tasks" highlight the tool's core function is removing the major challenges of the time tax and mental drain associated with writing. This finding confirms students choose to use AI as an immediate productivity booster, prioritizing instant efficiency over the necessary intellectual effort of independent thought with academic writing. The perceived risk of not using AI is framed as being "time consuming" and experiencing "cognitive exhaustion," solidifying the students' conscious choice to trade intellectual effort for immediate efficiency. Furthermore, beyond just being fast, Gen AI is also used by students as a powerful tool to improve their language and guarantee high quality in their academic writing. Students rely on the technology to make sure their final tasks and assignments meet high school standards, seeing the output as a perfect model for structure and sophisticated vocabulary.

This use is shown by reports of achieving a "perfect performance in making essay" and producing "better output" according to Discussant 3. Moreover, Discussant 2 & 3 noted a significant added benefit: they use the polished AI text to "learn new words" highlighting Gen AI's unexpected role in solving vocabulary gaps. Overall, the perceived challenge of traditional methods is the demand for high effort and time input, which often still results in imperfect output due to skill gaps. Students acknowledge that a potential benefit of not using AI is the ability to "useour critical thinking and our own understanding," but they frame the risk as it being "time consuming." Another student noted that without AI, the benefit is the chance to "practice my comprehension in writing especially in an essay." Gen AI's ability to allow for the strategic "offloading" of this high cognitive and temporal cost makes it the immediate, necessary solution in the students' view.

GenAI affects academic critical thought

In contrast to the benefits, the second major theme, Gen AI's impact on critical thought. Gen AI offers immediate efficiency by allowing students to outsource the intellectual effort required for critical thinking in academic writing. Furthermore, students themselves observed that their own critical thinking in academic writing were declining due to their heavy reliance on the AI tool. This ease of use, in turn, fosters strong dependence, which prevents the necessary challenge required for skill growth. Thus, resulting negative cycle stops learning by removing the struggle essential for developing expertise. Students connected their reliance on the tool to a slowdown in their personal development, Discussant 1 states: "I rely so much in Generative AI and my critical thinking is not improving". This finding confirms the idea that Gen AI is acting as a simple replacement for the student's own thinking. By removing the necessary intellectual friction and challenge, the technology prevents the student's mind from engaging in the genuine growth required to master critical thinking in academic writing. Moreover, the AI's ease of use makes it simple for students to hand off the entire intellectual task of academic writing. The Discussant 2 candid admission to answer the essay "because of AI" confirms that Gen AI is seen as a way to skip the core assignment. This validates the identified risk that students will simply "Turn off your brain"

and avoid critical effort. This finding addresses the specific choice students make when faced with complex academic tasks that require high effort. Whenever students encounter difficulties in applying their critical thinking in academic writing, Gen AI becomes their default solution and analytical substitute. By using the AI instead of struggling through the problem, students bypass essential skill practice and prevent themselves from improving their critical thinking abilities.

This avoidance is rooted in the fact that students identified several major difficulties that stop them from applying critical thinking during academic writing. These obstacles include procedural problems, such as not knowing how to "structure and write essay," and psychological problems like "mental blocking." Furthermore, Discussant 1, 2 also noted an internal cognitive barrier: the perceived slowness of their own thinking, which all push them toward the instant efficiency of AI. To address these difficulties, students reported using AI for the most difficult, core critical thinking tasks, confirming they are not just using it for grammar checks. Discussant 2, admitted to using the tool to "identify and analyze arguments", a function that requires high-order cognitive processes. This confirms that the AI is substituting for the complex analytical thought the student is supposed to perform. As a direct result of this substitution, the learner is removed from the difficult, real skill practice needed for mastery and growth. This pattern reinforces the theme that outsourced intellectual effort leads to cognitive stagnation in critical thinking in academic writing. Students recognize the conflict between Gen AI's powerful efficiency and the necessary challenge required for intellectual resilience. Recognizing this tension, the investigation shifted its focus to understanding the pedagogical implications and necessary instructional adjustments for English language teachers. Furthermore, three teachers who were handling English subjects were asked to share insights about the pedagogical implications of integrating Generative AI on ELT instruction in the classroom. The integration of Generative AI in ELT instruction presents profound pedagogical implications, which are best classified into positive transformations and negative risks concerning students' critical thinking in academic writing. To begin, the analysis reveals the pedagogical implications of integrating Generative AI necessitates a fundamental shift in the teacher's role from a traditional lecturer to a critical evaluation coach. Specifically, educators must move beyond the simple delivery of content to focus on teaching students how to scrutinize and validate machine-generated text. This transition is crucial because many students tend to accept GenAI outputs as factually correct without investigating their logical consistency. By acting as a mentor, the instructor ensures that technology serves to support rather than replace the student's own analytical effort.

Furthermore, a successful classroom environment must adopt a dual-path strategy that balances GenAI-assisted tasks with sessions of independent writing. This approach is designed to prevent students from falling into an "efficiency trap" where they become over-reliant on software for their foundational ideas. As a result, teachers can more track a student's logical progress by alternating between manual drafting and GenAI-supported refinement. Also, maintaining a dedicated space for manual effort is essential for building the mental stamina required for long-term academic success. In addition to structural changes, educators should design specialized tasks that require students to analyze and critique GenAI-generated drafts. Rather than using GenAI for simple text production, these assignments force learners to engage with the material to identify errors or biases. As a result, students must be able to justify their decisions to keep or modify specific suggestions made by the software. This strategic task design transforms the technology from a convenient shortcut into a sophisticated tool for advanced critical inquiry.

Finally, the successful integration of Generative AI in the ELT classroom depends on a strategic shift from passive tool use to active critical engagement, establishment of clear ethical guidelines and a roadmap for digital conduct. Specifically, this requires educators to serve as evaluation coaches while designing tasks that challenge students to critique machine-generated outputs. By balancing these digital tools with manual writing, schools can safeguard the development of independent critical thought. Moreover, standardized institutional policies are required to protect academic integrity and ensure that every student maintains their original voice. When these rules are in place, Generative AI serves as a powerful learning asset that empowers students rather than creating a mechanism for dependency. Thus, these pedagogical changes transform GenAI into a collaborative partner that strengthens, rather than replaces, the authentic academic voice.

DISCUSSION

The research sought to determine the effect of Gen-AI-Integrated ELT instruction on the students' critical thinking in academic writing. The results are presented and discussed by the specific research objectives, based on quantitative and thematic analysis of qualitative data. Based from the results, one instructional method positively influenced performance, the other failed to produce any measurable cognitive growth. Specifically, students in the experimental group struggled with the pre-test, characterizing their starting point as deficient. Despite the introduction of Generative AI throughout the intervention, the post-test data revealed a disappointing lack of progress among these learners. The majority of these learners remained stuck in the lowest scoring range, indicating that technology alone was insufficient to boost their critical thinking necessary for academic writing. Therefore, the finding suggests that an overreliance on GenAI was ineffective in fostering critical thinking in academic writing. This lack of growth emphasizes that without critical scaffolding, AI can become a barrier to genuine intellectual engagement in the students' critical thinking in academic writing rather than a helpful tool.

This finding aligns with the cognitive debt evidenced by neural deficits that highlights a corresponding challenge in assessment, explored in depth by Lee and Chen (2024) in their work, "Assessing Higher-Order Thinking in the Age of Generative AI." Their content analysis of student essays found that AI-generated texts often mask a lack of genuine critical thought behind fluent and sophisticated language mechanics. This methodological difficulty arises because traditional grading rubrics tend to reward surface-level linguistic features, which the LLM excels at producing, over the underlying analytical depth contributed by the student. In contrast, the control group utilizing traditional instruction started at the same very low performance level as their peers in the experimental group. This identical starting point was crucial because it established a fair and balanced baseline for evaluating the effectiveness of the conventional teaching

method. Following the completion of the instructional period, the post-test results showcased a significant and positive shift in students' performance in critical thinking in academic writing. This evidence confirms that the traditional ELT approach was successful in elevating students' critical thinking in academic writing through manual effort and human-led instruction. The collected data were further analyzed using T-tests to identify the significant differences between methods. Specifically, the results for the control group offer definitive proof that traditional instruction enhanced critical thinking in academic writing. By comparing pre-test and post-test scores, the data reveals a clear and positive growth in the students' critical thinking to handle academic writing tasks. This substantial improvement indicates that conventional teaching methods engaged students in the mental processes required for high-level logic. Unlike their peers, these students relied on human-led instruction to build their argumentative and analytical foundations through sustained manual effort. As a result, this group established a baseline for success that highlights the inherent effectiveness of active, manual learning over automated assistance using GenAI.

In contrast to these positive gains, the experimental group also showed a statistically significant difference in their scores, but this change was characterized by a notable decline. The post-test results dropped compared to the initial baseline, suggesting that the integration of Generative AI negatively impacted student performance. This negative trend implies that an overreliance on GenAI can compromise a student's critical thinking in academic writing process. Instead of acting as a helpful support system, the technology appeared to discourage the active engagement necessary for deep reflection and self-correction. As a result, students in this group experienced a measurable regression in their writing quality rather than the expected academic growth. These finding serves as a strong warning that the lack of a structured pedagogical framework in integrating GenAI may lead to cognitive stagnation. Therefore, educators must prioritize careful scaffolding and ethical guidance to ensure that AI supports rather than hinders long-term learning outcomes. Beyond these individual group trajectories, a final comparison between the two cohorts validates the findings by highlighting a significant performance gap at the end of the intervention. At the beginning of the research, statistical analysis confirmed that there was no significant difference in pre-test scores, ensuring both groups began with identical baselines. This comparable starting point ensures that the substantial difference in post-test results is related to the different instructional methods used. The post-test data revealed that the control group outperformed the experimental group by a significant statistical margin. This divide indicates that the GenAI-integrated group failed to keep pace with the intellectual progress achieved in the traditional setting. These results underscore the urgent necessity for a strategic pedagogical framework when introducing GenAI into the curriculum.

The interview themes emphasize the perceived critical thinking challenges faced by students in academic writing when utilizing Generative AI compared to those who solely rely on conventional methods. Students using traditional method often perceive as an exhausting and draining process that requires excessive time, whereas they view GenAI as an effortless shortcut. As a result, they utilize this tool as a primary mechanism to bypass linguistic barriers and achieve immediate "speed and perfect performance" in their essay writing tasks. In

this way, the strategic offloading of effort allows learners to meet high academic standards without engaging in the difficult and time-consuming labor of drafting from scratch. However, a major challenge arises because this heavy dependence on technology leads to a noticeable decline in students' critical thinking in academic writing. Specifically, by removing the "intellectual friction" of writing, the GenAI stunts long-term mental growth and causes students to "turn off their brains" during core analytical tasks. As a result, the tool's capacity for rapid output often outweighs the value of independent exertion, leading to intellectual stagnation instead of the mastery of critical thinking in academic writing.

Furthermore, relying on GenAI as a default solution creates a significant "cognitive debt" that hides a genuine lack of original reflection and ownership in a student's work. While conventional methods force students to remain the final authority on their own critical thinking necessary for academic writing, GenAI removes the productive struggle necessary for developing true expertise. Thus, this creates a paradox where a tool designed for efficiency becomes a primary obstacle to genuine, deep learning in students' critical thinking in academic writing. The perceived challenges center on a trade-off: students avoid the "time tax" and exhaustion of conventional methods, but they gain the challenge of "cognitive debt" and dependency that prevents students from mastering their critical thinking in academic writing. The synthesis of these findings highlights that the integration of Gen AI in ELT instruction presents this transition that is critically undermined by the risk of cognitive debt and skill atrophy. The efficiency allows the teacher to shift from a content expert to a critical evaluation coach, designing "AI-required" tasks based on a critical-use taxonomy to compel students toward higher-level analysis, thereby focusing on depth over mechanics. However, this positive shift is challenged by the negative risk that student over-reliance leads to the erosion of critical thinking in academic writing, as GenAI-generated fluency masks a genuine deficit in reflection and original thought.

Furthermore, the implication is that structured, independent essay drafting, without automated scaffolds, compels students to build the crucial analytic foundation that Gen AI may otherwise bypass. Therefore, the traditional method serves as the vital pedagogical benchmark for measuring the cognitive debt or enhancement provided by GenAI integration. The proposed teaching framework for the responsible and effective GEN-AI-Integrated ELT Instruction is the Critical-Scaffolding (CS) Framework. This incorporates several core elements: The Department of Education's Most Essential Learning Competencies (MELCs), English language teachers (ELTs) in the DepEd senior high school,principles, with theoretical foundation, senior high school learners and instructional phases that enable teachers to monitor progress and intervene whenever the technology begins to substitute for the student's original thought. This balanced approach ensures that Generative AI serves as a catalyst for discovery while promoting the student's critical thinking in academic writing.

Conclusion

This study proposed the Gen-AI Integrated Instruction to be used as one of the methods in the field of English Language Teaching. The intervention resulted in divergent learning trajectories: the traditional group demonstrated substantial

gains, while the Gen-AI–integrated group exhibited a significant decline in post-test performance. The negative results are explained by the fundamental pedagogical conflict revealed in the qualitative data, where the tool's immediate utility for high speed and low effort undermines intellectual exertion. Students recognize that this efficiency allows them to offload cognitive effort and bypass the essential struggle required for mastery, leading to a self-recognized risk of intellectual atrophy and dependence. The study's central finding is that achieving any genuine cognitive benefit is contingent upon the structured integration of the technology. This mandate requires adopting a dual-path cognitive process, balancing AI-assisted efficiency with deliberate, human-enforced tasks demanding maximum intellectual engagement. The implication is clear: without required cognitive challenge built into the task design, AI will substitute for, not supplement, critical thinking in academic writing. This occurs because when students are given tasks that can be finished by Generative AI specifically ChatGPT, the tool takes over the entire intellectual process in academic writing. This means students are delegating the cognitive work of analysis and synthesis to the AI, rather than using their own brains to develop complex arguments. As a result, for Gen AI to supplement and aid learning, English language teachers must design lessons that introduce cognitive challenge. This challenge forces students to engage with and evaluate the AI's output, rather than just accepting its suggestions. The success of GenAI integration, therefore, rests on the teacher's ability to evaluation coach who designs the lesson with "AI-required" tasks. Future ELT efforts must prioritize evidence-based Blended Learning Approaches that introduce mandatory cognitive challenge to ensure the technology supports, rather than substitutes for, the student's independent critical thinking in academic writing. The core takeaway is that the promise of AI integration cannot be realized without comprehensive and targeted training of English language teachers on effective GenAI-Integrated lesson planning and critical task design. Create instructional plans that make independent, critical engagement non-negotiable. In conclusion, the findings establish that the fundamental shift demands transitioning the ELT teacher's role from content expert to a critical

Recommendations

Based on the findings of the study, the researcher recommends the following:

1. The inevitability of technological integration mandates that English language teachers should explore and evaluate instructional models and alternative pedagogical interventions to identify the optimal balance between technological support and active, inquiry-based learning.
2. Teachers in ELT Instruction should develop evidence-based frameworks that maximize the tool's scaffolding benefits while preserving and strengthening students' independent critical thinking in academic writing.
3. English language teachers should act as instructional architects and critical coaches in their critical thinking in academic writing instruction.
4. Administrators in the institutions must develop and enforce clear policies that address the ethical and cognitive risks, specifically the threats of cognitive debt and compromised academic authenticity.
5. Policy Makers should implement immediate adaptation of assessment criteria to focus on verifiable student-led

critical activity, such as evaluating the process of revision or the quality of prompt refinement.
6. Policy Makers must provide an urgent mandate to close the reported gaps in teacher preparedness by implementing comprehensive, targeted professional development focused on strategic pedagogical competency and instructional risk mitigation.

REFERENCES

Alghamdi, F. M. (2023). Designing prompts for deep learning: A taxonomy for integrating ChatGPT to enhance critical thinking in ESL writing. TESOL Quarterly, 57(4), 1011–1028. (DOI not provided)

Andrews, R. (1995). Teaching and learning argument. Cassell.

Andrews, R. (2000). Introduction: Learning to argue in higher education. In R. Andrews & S. Mitchell (Eds.), Learning to argue in higher education (pp. 1–11). Boynton/Cook.

Bonnett, A. (2001). How to argue: A student's guide. Pearson Education.

British Council. (2024, July). Artificial intelligence and English language teaching: Preparing for the future (2nd ed.). https://www.teachingenglish.org.uk/sites/teacheng/files/AI_and_ELT_Jul_2024.pdf

Cooperman, S. R., & Brandão, R. A. (2024). AI tools vs AI text: Detecting AI-generated writing in foot and ankle surgery. Foot & Ankle Surgery: Techniques, Reports & Cases, Article 100367. https://doi.org/10.1016/j.fastrc.2024.100367

Creswell, J. W., & Plano Clark, V. L. (2014). Designing and conducting mixed methods research (2nd ed.). Sage.

Crompton, H. (2024). AI and English language teaching: Affordances and challenges. British Journal of Educational Technology. https://doi.org/10.1111/bjet.13460

Dela Cruz, R. S. (2012, November 28). The lack of critical thinking in Philippine schools. Manila Bulletin. https://ph.news.yahoo.com/lack-critical-thinking-philippine-schools-033314778.html

Edmett, A. (2023). Artificial intelligence and English language teaching – Report. British Council. (Use the 2024 2nd-ed. report above for a stable link.)

Elander, J., Harrington, K., Norton, L., Robinson, H., & Reddy, P. (2006). Complex skills and academic writing: A review of evidenceabout the types of learning required to meet core assessment criteria. Assessment & Evaluation in Higher Education, 31(1), 71–90. https://doi.org/10.1080/02602930500262494

Essien, A., Bukoye, O. T., O'Dea, C., &Kremantzis, M. (2024). The influence of AI text generators on critical thinking skills in UK business schools. Studiesin Higher Education. https://doi.org/10.1080/03075079.2024.2316881

Hyland, K. (2014). English for academic purposes. In C. Leung & B. V. Street (Eds.), The Routledge companion to English studies (pp. 392–404). Routledge. https://doi.org/10.4324/9781315852515

Irfan,N., & Robbins, S. (2024). Generative AI in education: Teaching strategies for cognitive skills and critical thinking with ChatGPT. Research Gate. https://doi.org/10.13140/RG.2.2.14469.15844

Kojima, T., Gu, S. S., Reid, M., Matsuo, Y., & Iwasawa, Y. (2022). Large language models are zero-shot reasoners arXiv. https://doi.org/10.48550/arXiv.2205.11916

Kos'myna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I., & Maes, P.

(2025). Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task. arXiv preprintarXiv:2506.08872

Lee, S. K., & Chen, R. (2024). Assessing higher-order thinking in the age of generative AI: A content analysis of student essays. Journal of Educational Psychology, 98(3), 450–465. (DOI not provided)

Liu, J., et al. (2025). How do generative artificial intelligence (AI) tools and large language models influence self-regulated learning and critical thinking skills? Smart Learning Environments. https://doi.org/10.1186/s40561-025-00406-0

Lugtu, R. C., Jr. (2019, February 7). Pandemic ignorance. The Manila Times. https://www.manilatimes.net/pandemic-ignorance/507946

Major, K., & Chiarelott. (2023). Slow down: Generative AI, faculty reactions, and the role of critical thinking in writing instruction. Double Helix, 11, Article 3. https://doi.org/10.37514/DBH-J.2023.11.1.03

Nori, H., King, N., McKinney, S. M., Carignan, D., & Horvitz, E. (2023). Capabilities of GPT-4 on medical challengeproblems. arXiv. https://doi.org/10.48550/arXiv.2303.13375

Schmohl, T., Watanabe, A., Fröhlich, N., & Herzberg, D. (2020). How can artificial intelligence improve the academic writing of students? In Conference proceedings: The Future of Education (10th ed., pp. 168171). Filodiritto Editore. https://www.pedocs.de/volltexte/2023/27913/pdf/Schmohl_et_al_2020_How_can_artificial_intelligence.pdf

Shanto, S., & Ahmed, Z. (2024). Enriching the learning process with generative AI: A proposed framework to cultivate critical thinking in higher education using ChatGPT. Tuijin Jishu/Journal of Propulsion Technology, 45(1),10014055. https://doi.org/10.52783/tjjpt.v45.i01.4680

Swales, J. M., & Feak, C. (2012). Academic writing for graduate students: Essential tasks and skills (3rd ed.).University of Michigan Press. https://doi.org/10.3998/mpub.2173936

Tapper, J. (2004). Student perceptions of how critical thinking is embedded in a degree program. Higher Education Research & Development, 23(2), 199–222. (DOI not provided)

Tashakkori, A., & Creswell, J. W. (2007). The new era of mixed methods. In Handbook of research on teaching (2nd ed., pp. 249–264). Blackwell. (Link/DOI not provided)

Tsui, L. (2002). Fostering critical thinking through effective pedagogy. The Journal of Higher Education, 73(6), 740–763. (DOI not provided)

Wardle, E. (2025, March 30). Writing is thinking. The Hechinger Report. https://hechingerreport.org/writing-is-thinking/

Weidmann, A. E. (2024). Artificial intelligence in academic writing and clinical pharmacy education: Consequences and opportunities. International Journal of Clinical Pharmacy. https://doi.org/10.1007/s11096-024-01705-1

****