ENHANCING CRITICAL THINKING AND WRITING QUALITY THROUGH PROMPT ENGINEERING: A STUDY OF HIGH SCHOOL STUDENTS AT SMAN 1 DAWARBLANDONG

Nur Kholis Majid | English Language Education, Universitas Islam Majapahit, Mojokerto, Indonesia
Nuermajid144@gmail.com

Abstrak

This study investigated the effect of prompt engineering on the critical thinking (CT) skills and writing quality (WQ) of eleventh-grade students at SMAN 1 Dawarblandong. The research employed a quasi-experimental pre-test and post-test control group design involving 64 participants, consisting of 32 students in the experimental group (Class XI-3) and 32 students in the control group (Class XI-4). Students in the experimental group participated in six sessions of prompt engineering based on the TCI-RATEL framework, whereas those in the control group received conventional writing instruction without structured guidance in prompt formulation. The collected data were analysed using the independent samples t-test. The findings revealed that prompt engineering had a significant positive effect on students' critical thinking skills, $t(62) = 4.743$, $p < .001$, with a large effect size (Cohen's $d = 1.19$). Likewise, a significant improvement was found in students' writing quality, $t(62) = 5.406$, $p < .001$, with a large effect size (Cohen's $d = 1.35$). This study is urgent given Indonesia's persistently low PISA 2022 performance in analytical reasoning and critical evaluation, coupled with secondary students' documented tendency to use generative AI passively; empirical evidence on prompt engineering within Indonesian senior high school EFL classrooms remains scarce, leaving teachers without a validated, contextually grounded instructional approach. These findings indicate that integrating structured prompt engineering into EFL writing instruction is effective in enhancing both critical thinking and writing quality among secondary school students.

Keywords: prompt engineering; critical thinking; writing quality; TCI-RATEL framework; EFL writing.

Introduction

The rapid development of generative artificial intelligence (AI) has significantly transformed English as a Foreign Language (EFL) writing instruction by providing students with various forms of learning support. AI-powered applications such as ChatGPT enable learners to generate ideas, improve sentence structure, and correct grammatical errors in a short time, offering assistance that was previously difficult to obtain, particularly for secondary school students. Although these technologies create valuable learning opportunities, they also introduce new instructional challenges

concerning how students interact with AI during the writing process. During the researcher's Teaching Practicum Programme (Praktik Pengalaman Lapangan or PPL) at SMAN 1 Dawarblandong, it was observed that many students relied heavily on ChatGPT when completing English writing assignments. Rather than using AI as a learning aid, students frequently accepted AI-generated responses without evaluating their accuracy, relevance, or appropriateness. Although the integration of AI into education is not inherently problematic, the passive manner in which students utilise these tools presents important pedagogical concerns.

Similar patterns have been documented beyond the research site. A survey of 1,024 upper secondary students in Norway found that most participants interacted with generative AI passively, treating it primarily as a shortcut for completing assignments rather than as a tool to support learning and reflection (Bueie et al., 2025). These observations suggest that the passive use of AI is not unique to SMAN 1 Dawarblandong but reflects a wider trend in AI-assisted writing instruction that warrants closer pedagogical attention.

These classroom observations also reflect a broader educational issue related to Indonesian students' higher-order thinking abilities, as Indonesia's performance in the PISA 2022 assessment continues to indicate relatively low literacy achievement, with many secondary school students still experiencing difficulties in analytical reasoning and critical evaluation (PISA 2022 as cited in Saifulloh et al., 2023). Within EFL learning specifically, writing represents one of the most demanding language skills because students are expected to integrate linguistic competence, including grammar and vocabulary, with rhetorical abilities such as organising ideas and constructing convincing arguments.

Despite the rapid growth of international research concerning AI-assisted writing and prompt engineering, studies investigating these topics within Indonesian secondary school contexts remain limited. Existing Indonesian research has begun to explore the educational use of AI in EFL classrooms (Ilma et al., 2020; Rizkiani, 2024); however, most studies have focused on university students rather than learners at the senior high school level. Similarly, much of the international literature on prompt engineering and creative writing has been conducted in higher education settings or outside Indonesia (Kabeer et al., 2025). Considering the distinctive characteristics of Indonesian education, including the implementation of the Merdeka Curriculum, its emphasis on developing 4C competencies, differences in technology access, students' literacy performance, and sociocultural learning contexts, there is a need for empirical research specifically examining the integration of AI within Indonesian secondary schools.

Based on these considerations, this study investigates the effect of structured prompt engineering using the TCI-RATEL framework on the critical thinking skills and writing quality of eleventh-grade EFL students. Accordingly, this study seeks to answer two research questions: (1) How does prompt engineering instruction affect the critical thinking skills of eleventh-grade students at SMAN 1 Dawarblandong in their English writing assignments? (2) How does prompt engineering instruction affect the writing quality of eleventh-grade students at SMAN 1 Dawarblandong? It is hypothesised that students receiving structured prompt engineering instruction will demonstrate significantly better critical thinking skills and writing quality than those receiving conventional instruction with unstructured AI use. Ultimately, this study is expected to provide empirical evidence regarding the relationship between prompt engineering, critical thinking, and writing quality in Indonesian secondary education while offering practical instructional recommendations that support the implementation of the Merdeka Curriculum, particularly its emphasis on developing 21st century competencies through the 4C framework.

Literature Review

Several studies underscore the importance of critical engagement when students use generative AI for writing. Rizkiani (2024) reported that students in AI-assisted writing environments frequently accepted AI-generated information without sufficiently questioning or evaluating its quality, using a critical thinking framework that emphasises clarity, logical reasoning, argument development, and evidence integration; the study found that students experienced considerable difficulty constructing clear thesis statements, organising ideas coherently, developing logical arguments, and synthesising information from multiple sources. Although AI applications such as ChatGPT and Grammarly were found to improve writing quality when used strategically, excessive dependence on AI-generated suggestions was also shown to reduce students' opportunities to think critically throughout the writing process (Rizkiani, 2024). Complementing these classroom-level findings, S. Tian and Xiang (2025) identified critical thinking as one of the strongest predictors of effective engagement with generative AI in writing tasks, showing that students with stronger critical thinking skills were more likely to analyse, evaluate, and revise AI-generated responses instead of accepting them uncritically, resulting in higher-quality writing outcomes. At a broader level, Liu et al. (2026) identified critical thinking disposition as an important predictor of students' academic achievement, demonstrating that stronger critical thinking indirectly improves academic performance through increased academic self-efficacy, while Li (2022) argued that students possessing stronger critical thinking skills tend to produce writing containing more coherent arguments, stronger reasoning, and more effective use of supporting evidence.

Evidence from recent scholarship further supports the pedagogical value of teaching students to engineer prompts deliberately rather than allowing unstructured AI use. Walter (2024) argues that prompt engineering can function as an effective instructional strategy because developing effective prompts requires learners to analyse tasks, evaluate AI responses, and continuously revise their prompts to obtain more accurate and relevant outputs. In a similar vein, Federiakin et al. (2024) describe prompt engineering as an emerging 21st century competence that promotes students' cognitive engagement with digital technologies beyond simple technical operation. At the secondary level, Kabeer et al. (2025) found that a structured, phased prompt engineering intervention enhanced Indonesian secondary school students' creative writing skills, lending further support to the claim that explicit instruction in prompt formulation, rather than incidental exposure to AI tools, is what drives measurable learning gains.

Building on these perspectives, this study adopts the TCI-RATEL framework as its conceptual foundation for teaching prompt engineering. The framework organises prompt construction into two hierarchical stages: the component level, consisting of Task, Context, and Instruction (TCI), and the element level, comprising Role, Audience, Tone, Examples, and Limits (RATEL). The component level establishes the core purpose and boundaries of a prompt, while the element level refines it with contextual detail that shapes the tone, perspective, and specificity of the AI-generated response. Within EFL writing instruction, this two-stage structure gives students an explicit, teachable procedure for translating a writing task into a prompt that elicits a more useful, purpose-driven output, rather than relying on trial-and-error prompting.

Taken together, the reviewed literature establishes two convergent lines of argument: first, that unstructured AI use in writing tends to encourage passive, uncritical engagement (Bueie et al., 2025; Rizkiani, 2024); and second, that critical thinking and structured prompt formulation are closely interrelated competencies that can be deliberately cultivated through instruction (Federiakin et al., 2024; S. Tian & Xiang, 2025; Walter, 2024). However, most of the empirical evidence linking prompt engineering instruction to measurable gains in both critical thinking and writing quality

comes from higher education or non-Indonesian secondary contexts (Kabeer et al., 2025). This gap motivates the present study's quasi-experimental investigation of the TCI-RATEL framework among Indonesian senior high school EFL students.

Methodology

This study employed a quasi-experimental research design using a pre-test post-test control group design to examine the effect of prompt engineering instruction on students' critical thinking skills and writing quality. This research design was selected because it allows the researcher to compare learning outcomes between an experimental group and a control group while measuring changes that occur before and after the intervention. The use of a quasi-experimental approach was considered appropriate, as random assignment of individual students was not feasible due to the existing classroom organisation at the research site (Saifulloh, Retnaningdyah, & Mustofa, 2025c)

During the intervention, the experimental group participated in English writing instruction that incorporated structured prompt engineering activities, whereas the control group received conventional writing instruction and used AI tools without explicit guidance on prompt construction. Prior to and following the instructional treatment, both groups completed a pre-test and post-test writing assessment using a validated research instrument. The comparison of the assessment results enabled the researcher to determine the extent to which the prompt engineering instruction contributed to improvements in students' critical thinking skills and writing quality while accounting for any initial differences between the two groups (Zhang et al., 2025).

Population and Sample

The participants in this study were drawn from Class XI-3 and Class XI-4 of SMAN 1 Dawarblandong, a public senior high school located in Mojokerto Regency, East Java, Indonesia, during the 2025/2026 academic year. Eleventh-grade students were chosen because they possess an adequate level of English proficiency to complete narrative writing tasks effectively while remaining in an important stage of developing higher-order thinking and AI literacy competencies.

The sample was determined using purposive sampling, with two intact classes selected according to three predetermined criteria: (a) comparable levels of English proficiency based on students' academic performance in the previous semester, (b) similar patterns of AI tool usage observed during the researcher's teaching practicum, and (c) willingness and availability to participate throughout the three-week intervention. Class XI-3, consisting of 32 students, served as the experimental group, whereas Class XI-4, also comprising 32 students, functioned as the control group. Before the implementation of the intervention, the equivalence of the two groups was examined statistically through an Independent Samples t-test using the pre-test scores. This procedure was conducted to ensure that both groups had comparable initial abilities, thereby satisfying the methodological requirements of quasi-experimental research and supporting the validity of subsequent comparisons (Saifulloh & Anam, 2022)

Variable and Instrument

The study involved one independent variable and two dependent variables. The independent variable was prompt engineering instruction, which consisted of a structured instructional intervention designed to train students to formulate clear, specific, and purpose-driven prompts for ChatGPT using the TCI-RATEL framework to support English writing activities.

The first dependent variable was critical thinking (CT) skills, which were assessed using the analytical rubric developed by W. Tian et al. (2026). This rubric measures three key dimensions of critical thinking: Analysing Arguments, Developing Arguments, and Evaluating Perspectives. Each dimension is rated on a four-point scale ranging from 1 (Beginning) to 4 (Exemplary), resulting in a maximum overall score of 12 points.

The second dependent variable was writing quality (WQ), which was evaluated using the analytical rubric proposed by Diego and Protacio (2025). The rubric assesses students' writing across five weighted components: Content (30%), Language Use (25%), Organisation (20%), Vocabulary (20%), Mechanics (5%), with a maximum possible score of 100 points.

Before data collection, both assessment rubrics were reviewed by two experts in English as a Foreign Language (EFL) and educational assessment to establish their content validity. Revisions were incorporated based on the reviewers' recommendations to ensure that each instrument accurately represented the constructs being measured. To evaluate scoring consistency, inter-rater reliability between the two assessors was calculated using Cohen's Kappa, with a minimum acceptable value of $\varkappa \geq .80$ (Karaer et al., 2024). In addition, the internal consistency of the rubric scores was examined using Cronbach's Alpha, with $\alpha \geq .70$ considered the acceptable threshold for reliability.

Procedure

The study was implemented over a period of three consecutive weeks and consisted of three sequential phases. During the pre-intervention phase, the researcher began by reviewing the fundamental concepts of narrative writing, including its social function, generic structure, and essential language features. After this brief review, both the experimental and control groups completed a pre-test by writing a narrative text based on an Indonesian folktale. At this stage, students were not provided with any instruction regarding prompt engineering, allowing the pre-test to measure their initial levels of critical thinking and writing quality.

The intervention phase was conducted exclusively with the experimental group through the implementation of the TCI-RATEL prompt engineering framework. The researcher first explained the concept of prompt engineering, its role in AI-assisted writing, and the characteristics of effective and ineffective prompts. Students were then guided to formulate structured prompts that explicitly incorporated narrative elements and task requirements. Throughout the instructional activities, they were encouraged to critically evaluate AI-generated responses and improve their prompts through an iterative refinement process. In contrast, the control group participated in conventional English writing instruction during the same period without receiving explicit training in prompt engineering strategies.

In the post-intervention phase, both groups completed a post-test using a narrative writing task based on the same Indonesian folktale theme. Students in the experimental group applied the prompt engineering techniques they had learned during the intervention, whereas those in the control group completed the writing task without structured prompt guidance. The completed writing products from both groups were subsequently collected and assessed to examine the impact of the instructional intervention on students' critical thinking skills and writing quality.

Data Analysis

Quantitative data were analysed using IBM SPSS Statistics through a series of systematic procedures. The analysis began with a preliminary assumption testing, in which the normality of the gain score distributions was examined using the Shapiro-Wilk test, as the test is appropriate for sample sizes of fewer than 32 participants per group. The homogeneity of variances between the experimental and control groups was subsequently evaluated using Levene's Test.

After confirming that both statistical assumptions had been satisfied, the Independent Sample t-test was conducted to examine the two research hypotheses by comparing the gain scores of the experimental and control groups for both Critical Thinking (CT) Skills and Writing Quality (WQ). Finally, Cohen's d was calculated to estimate the magnitude of the intervention effect and determine its practical significance. The effect size was interpreted according to Cohen's (1988) conventional criteria, namely small ($d \geq 0.20$), medium ($d \geq 0.50$), and large ($d \geq 0.80$).

Findings and Discussion

This section presents the results of the study and discusses their implications. It begins with the descriptive statistics of the pre-test and post-test scores for both the experimental and control groups, followed by the results of the assumption testing (normality and homogeneity of variance) that justified the use of parametric statistics. The Independent Samples t-test and Cohen's d results are then reported to answer the two research questions concerning the effect of prompt engineering instruction on students' critical thinking skills and writing quality. Building on these statistical findings, the discussion examines the effect of prompt engineering on critical thinking, its effect on writing quality, and the interrelationship between the two constructs, situating the results within existing theory and prior empirical studies. Finally, the section considers the broader relevance of these findings to the implementation of Deep Learning principles and the 4C competencies within Indonesia's Merdeka Curriculum.

Descriptive Statistics

Table 1 provides an overview of the descriptive statistics for the experimental and control groups at both the pre-test and post-test stages. The similarity of the pre-test mean scores for Critical Thinking (CT) and Writing Quality (WQ) indicates that the two groups were comparable before the intervention. After the implementation of the instructional treatment, the experimental group showed markedly greater improvements than the control group across both variables. Specifically, the experimental group achieved a mean gain that was 8.53 points higher in Critical Thinking and 9.19 points higher in Writing Quality, suggesting a stronger positive effect of the prompt engineering instruction.

Table 1. Descriptive Statistics of Pre-Test and Post-Test Scores

Statistic	Pre CT Exp	Pre CT Ctrl	Pre WQ Exp	Pre WQ Ctrl	Post CT Exp	Post CT Ctrl	Post WQ Exp	Post WQ Ctrl
Mean	58.25	58.59	63.84	63.09	75.66	67.03	81.47	73.19
CI Lower	56.80	57.06	62.42	61.71	73.07	63.89	78.40	70.59
CI Upper	59.70	60.12	65.27	64.47	78.24	70.18	84.53	75.79
Trimmed Mean	58.12	58.49	63.86	63.03	75.76	67.21	81.60	73.22
Median	58.00	58.00	63.00	63.50	76.00	68.50	82.00	74.00
Variance	16.194	17.991	15.684	14.668	51.523	76.096	72.322	51.964
Std. Deviation	4.024	4.242	3.960	3.830	7.178	8.723	8.504	7.209
Minimum	52	52	56	56	57	49	61	59
Maximum	68	67	71	72	91	83	98	86
Range	16	15	15	16	34	34	37	27
Interquartile Range	6	7	8	6	9	14	7	12
Skewness	0.388	0.361	-0.016	0.078	-0.414	-0.436	-0.214	-0.004
Kurtosis	-0.442	-0.629	-0.894	0.032	0.454	-0.566	0.585	-0.835

Note: Eks = experimental group; Kon = control group.

Assumption Testing and Hypothesis Testing

Before testing the research hypotheses, the assumptions underlying the use of parametric statistics were examined. The normality of the pre-test and post-test scores for both dependent variables were assessed using the Shapiro-Wilk test. All significance values were greater than the .05 threshold, indicating that the data were normally distributed. Specifically, the pre-test Critical Thinking (CT) scores showed normal distributions for both the experimental group ($W = .963$, $p = .322$) and the control group ($W = .960$, $p = .276$). Likewise, the pre-test Writing Quality (WQ) scores met the normality assumption in the experimental group ($W = .969$, $p = .462$) and the control group ($W = .974$, $p = .610$). Similar results were obtained for the post-test scores, with CT showing $W = .964$, $p = .356$ for the experimental group and $W = .962$, $p = .310$ for the control group, while WQ yielded $W = .951$, $p = .154$ and $W = .976$, $p = .691$, respectively.

The homogeneity of variance assumption was subsequently evaluated using Levene's Test on the gain scores. The results indicated that the variances of both dependent variables were homogeneous across the two groups. For Critical Thinking, Levene's Test produced F = 1.217, p = .274, whereas for Writing Quality, the test yielded F = .006, p = .937. Because both significance values exceeded .05, the assumption of equal variances was satisfied. Accordingly, the Equal Variances Assumed row was used in interpreting the results of the Independent Samples t-test.

Table 2. Independent Samples Test on Gain Scores

Variable	Assumption	Levene F	Levene Sig.	t	df	Sig. (2-tailed)	Mean Diff.	SE Diff.	95% CI Lower	95% CI Upper
gain_ct	Equal variances assumed	1.217	.274	4.743	62	.000	8.531	1.799	4.936	12.127
gain_ct	Equal variances not assumed			4.743	60.785	.000	8.531	1.799	4.934	12.128
gain_wq	Equal variances assumed	.006	.937	5.406	62	.000	9.188	1.700	5.790	12.585
gain_wq	Equal variances not assumed			5.406	61.462	.000	9.188	1.700	5.790	12.585

The analysis revealed a statistically significant difference in Critical Thinking (CT) gain scores between the two groups, t (62) = 4.743, p < .001, with a mean difference of 8.531 points (95% CI [4.936, 12.127]). Since the assumption of equal variances was satisfied (F = 1.217, p = .274), the results were interpreted using the Equal Variances Assumed output. Therefore, the null hypothesis (H_{10}) was rejected, indicating that prompt engineering instruction significantly improved students' critical thinking skills compared with conventional instruction.

A similar pattern was observed for Writing Quality (WQ). The Independent Samples t-test showed a statistically significant difference in gain scores between the experimental and control groups, t (62) = 5.406, p < .001, with a mean difference of 9.188 points (95% CI [5.790, 12.585]). As Levene's Test confirmed the homogeneity of variances (F = .006, p = .937), the Equal Variances Assumed results were adopted for interpretation. Consequently, the null hypothesis (H_{20}) was rejected, demonstrating that prompt engineering instruction had a significant positive effect on students' writing quality.

To complement the hypothesis testing, Cohen's d was calculated to estimate the magnitude of the intervention effects. The effect size for Critical Thinking was d = 1.19, while Writing Quality produced an effect size of d = 1.35. According to Cohen's (1988) guidelines, both values represent large effect sizes, indicating that the intervention was not only statistically significant but also educationally meaningful.

The Effect of Prompt Engineering Instruction on Critical Thinking Skills

The findings of this study demonstrate that prompt engineering instruction contributed significantly to the improvement of students' Critical Thinking (CT) skills. This result supports the theoretical perspective proposed by Walter (2024), who argues that prompt engineering can serve as an effective instructional strategy for fostering critical thinking when it is systematically integrated into classroom learning. Throughout the intervention, students were required to formulate structured prompts, critically examine AI-generated responses, and refine their prompts through an iterative process to obtain outputs that were more accurate and relevant. These learning activities encouraged students to engage in higher-order cognitive processes, including analyzing problems, evaluating information, and making informed decisions, which correspond to the critical thinking dimensions assessed in this study, namely analyzing arguments, developing reasoning, and evaluating perspectives as measured using the rubric developed by W. Tian et al. (2026).

The present findings also extend the work of S. Tian and Xiang (2025), who identified critical thinking as one of the strongest predictors of effective engagement with generative AI in writing tasks. While their study demonstrated that students with stronger critical thinking skills tend to interact with AI more strategically, the results of the current study suggest a complementary relationship. Specifically, structured prompt engineering instruction appears to enhance students' critical thinking by guiding them to interact with AI in a more reflective, analytical, and purposeful manner. This indicates that prompt engineering not only benefits from critical thinking but can also function as a means of cultivating it within the context of secondary-level EFL instruction.

Furthermore, the findings are consistent with the quasi-experimental study conducted by Baş and Gürsoy (2021), which reported that English language instruction explicitly incorporating critical thinking activities such as argument development, evidence evaluation, and reflective inquiry resulted in significantly greater improvements in students' critical thinking than conventional teaching approaches. This similarity reinforces the view that critical thinking does not develop automatically through routine classroom practices; instead, it requires carefully designed instructional strategies that deliberately engage students in higher-order cognitive processes.

The improvement in students' CT performance also has important implications for the Indonesian educational context. Indonesia's performance in PISA 2022 highlighted persistent challenges in students' higher-order thinking abilities, particularly in analytical reasoning and critical evaluation ("PISA 2022 Results," 2022). The significant gains observed in this study suggest that prompt engineering instruction may represent a practical and contextually relevant pedagogical approach for strengthening these competencies. By encouraging students to engage critically with AI-generated information rather than accepting it passively, prompt engineering can support the development of higher-order thinking skills that are increasingly emphasized within the Merdeka Curriculum and are essential for learning in the digital era(Saifulloh, 2021a, 2021b, 2026; Saifulloh et al., 2026; Saifulloh, Retnaningdyah, & Mustofa, 2025a, 2025b; Saifulloh, Retnaningdyah, Mustofa, et al., 2025).

The Effect on Writing Quality

The findings of this study indicate that prompt engineering instruction produced a substantial improvement in students' Writing Quality (WQ). This result is in line with the quasi-experimental study conducted by Diego and Protacio (2025), which reported that senior high school students who received structured AI-assisted writing instruction achieved significantly better writing performance across the dimensions of content, organization, vocabulary, language use, and mechanics than those who experienced conventional instruction. The present study extends these findings to the Indonesian EFL context, demonstrating that students who were explicitly trained to formulate effective prompts produced narrative texts with richer content, more coherent organization, more appropriate lexical choices, greater grammatical accuracy, and stronger mechanical correctness.

The findings also support those of Kabeer et al. (2025), who concluded that structured prompt engineering activities positively influence students' writing development. Collectively, these studies reinforce the argument that meaningful improvements in writing are more likely to occur when students receive explicit guidance in constructing prompts rather than relying solely on unrestricted access to AI-generated assistance.

Interestingly, the intervention produced a slightly larger effect on Writing Quality ($d = 1.35$) than on Critical Thinking ($d = 1.19$). The direct relationship between prompt construction and the quality of written output may explain this difference. Within the TCI-RATEL framework, students were guided to formulate prompts that explicitly specified narrative elements, including characters, setting, plot development, and writing style. Such structured guidance enabled students to use AI

as a learning scaffold that supported idea generation and text development, rather than merely serving as a shortcut for completing writing tasks.

These findings also address concerns raised by Bueie et al. (2025), who observed that many secondary school students interact with generative AI passively by accepting AI-generated responses without sufficient evaluation or revision. A similar pattern was identified during the researcher's teaching practicum and has also been reported by Rizkiani (2024). The clear difference in writing performance between the experimental and control groups suggests that explicit prompt engineering instruction can reduce passive dependence on AI by encouraging students to engage more critically and purposefully throughout the writing process.

The Interrelationship Between Critical Thinking and Writing Quality

The improvement in students' writing quality appears to be closely associated with the development of their critical thinking skills during the intervention. As students became more capable of analyzing, evaluating, and refining AI-generated responses, they were better able to select relevant information, organize ideas logically, and integrate those ideas into their own writing. Consequently, their written work demonstrated stronger content development, improved coherence, greater grammatical accuracy, and more appropriate vocabulary use.

This finding supports the theoretical perspectives proposed by Zou et al. (2022) and Khairuddin et al. (2024), who argue that higher-order thinking skills and writing performance are closely interconnected. Students who think critically are generally better equipped to construct logical arguments, organize information effectively, and communicate ideas with greater clarity. The present findings are further supported by the quasi-experimental study conducted by Zhang et al. (2025), which demonstrated that instructional strategies promoting collaborative reasoning significantly improved students' critical thinking abilities. Although Zhang et al. (2025) reported effect sizes ranging from 0.51 to 0.84, the current study produced a larger effect size ($d = 1.19$) for critical thinking, suggesting that structured prompt engineering may provide an equally or even more effective approach for promoting higher-order thinking within EFL writing instruction.

Relevance to Deep Learning and the Merdeka Curriculum

Beyond its contribution to AI-assisted language learning, this study also offers practical implications for the implementation of Pembelajaran Mendalam (Deep Learning) within Indonesia's Merdeka Curriculum. Deep Learning promotes student-centered learning experiences that encourage meaningful understanding, reflective thinking, and active engagement rather than superficial knowledge acquisition (Riani & Sujarwati, 2025). Its primary objective is to cultivate higher-order thinking skills through learning experiences that are meaningful, mindful, and enjoyable.

The instructional approach implemented through the TCI-RATEL framework reflects these core principles. Meaningful Learning was fostered by encouraging students to connect their prior understanding of narrative texts with the purposeful construction of AI prompts. Mindful Learning was promoted through continuous prompt revision and critical evaluation of AI-generated responses, requiring students to monitor and regulate their own thinking throughout the writing process. Meanwhile, Joyful Learning emerged from the creative nature of the narrative writing activities, which allowed students to explore ideas independently while interacting with AI in an engaging and motivating way (Luckita et al., 2025; Ambarwati & Wakhidah, 2025).

These findings suggest that prompt engineering instruction represents a practical strategy for translating the principles of Deep Learning into classroom practice. More importantly, it provides English teachers with a structured approach to integrating AI into writing instruction while simultaneously supporting the development of the 4C competencies, particularly critical thinking, creativity, communication, and collaboration, emphasized in the Merdeka Curriculum.

Conclusion

This final section summarises the main findings of the study in relation to the two research questions, outlines their theoretical and pedagogical implications, and offers suggestions for classroom practice and future research. It first restates the statistical evidence for the effect of prompt engineering instruction on students' critical thinking skills and writing quality, then discusses what these results mean for EFL teachers seeking to integrate AI tools into writing instruction under the Merdeka Curriculum, and closes with recommendations for future studies that could extend or strengthen the present findings.

The findings of this study demonstrate that the structured prompt engineering instruction based on the TCI-RATEL framework significantly enhanced both the critical thinking (CT) skills and writing quality (WQ) of eleventh-grade EFL students at SMAN 1 Dawarblandong. Statistical analysis confirmed significant differences between the experimental and control groups for both variables, with critical thinking showing $t(62) = 4.743$, $p < .001$, $d = 1.19$, and writing quality showing $t(62) = 5.406$, $p < .001$, $d = 1.35$. The large effect sizes indicate that the intervention was not only statistically significant but also educationally meaningful.

These findings show that explicit instruction in prompt engineering is considerably more effective than allowing students to use generative AI without structured guidance, enabling students to engage more critically with AI-generated information and produce higher-quality writing. The results also suggest that prompt engineering provides a practical means of implementing Deep Learning principles within English language classrooms while supporting the higher-order thinking skills promoted by the Merdeka Curriculum. English teachers are therefore encouraged to integrate structured prompt engineering instruction, such as the TCI-RATEL framework, into EFL writing lessons rather than restricting students' access to AI tools, and to guide learners in using these technologies critically and responsibly. Future studies are recommended to employ true experimental designs with random assignment, larger and more diverse samples, and longitudinal designs, complemented by qualitative approaches such as think-aloud protocols or classroom observations to gain deeper insight into how prompt engineering shapes students' cognitive processes and writing development.

References

Ambarwati, S., & Wakhidah, N. (2025). Analisis persepsi calon guru terhadap penerapan pembelajaran mendalam (Deep Learning) dalam pembelajaran IPA. Jurnal Pendidikan Tambusai, 3(12), 1–14. https://doi.org/10.62281/0t9wgv74

Baş, H. K., & Gürsoy, E. (2021). The effect of critical thinking embedded English course design to the improvement of critical thinking skills of secondary school learners. Thinking Skills and Creativity. https://doi.org/10.1016/j.tsc.2021.100910

Bueie, A. A., Skar, G. B., & Graham, S. (2025). High school students' use and beliefs about generative artificial intelligence and writing in school. Reading and Writing. https://doi.org/10.1007/s11145-025-10702-3

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

Diego, S. S., & Protacio, A. V. (2025). AI-integrated writing approach and senior high school students' writing competencies in English. International Journal of English Literature and Social Sciences. https://doi.org/10.22161/ijels.103.97

Federiakin, D., Molerov, D., Zlatkin-Troitschanskaia, O., Maur, A., & Reynolds, B. L. (2024). Prompt engineering as a new 21st century skill. Frontiers in Education. https://doi.org/10.3389/feduc.2024.1366434

Ilma, I. N., Susantini, E., & Purnomo, T. (2020). The development of science learning devices with guided inquiry model through flipbook media on human respiratory system material to train critical thinking skills of junior high school students. International Journal for Educational and Vocational Studies. https://doi.org/10.29103/ijevs.v2i5.2508

Kabeer, A., Bhat, R. A., Antony, S., & Tramboo, I. A. (2025). Enhancing creative writing skills in secondary school students through prompt engineering and artificial intelligence. Forum for Linguistic Studies, 7(3), 800–815. https://doi.org/10.30564/fls.v7i3.8511

Karaer, G., Hand, B., & French, B. F. (2024). Examining the impact of science writing heuristic (SWH) approach on development of critical thinking, science and language skills of students with and without disabilities. Thinking Skills and Creativity. https://doi.org/10.1016/j.tsc.2023.101443

Khairuddin, Z., Daud, K. M., Anuar, N., Satimin, O., Mohd Yusof, F. H., & Sabri, S. (2024). Relationship between perceived students' critical thinking skills and academic writing skills. International Journal of Research and Innovation in Social Science, 8(3s), 4709–4721. https://doi.org/10.47772/IJRISS.2024.803343S

Li, S. (2022). Constructing a critical thinking evaluation framework for college students majoring in the humanities. Frontiers in Psychology. https://doi.org/10.3389/fpsyg.2022.1017885

Liu, W., Ma, H., Ma, S., Chen, Y., & Li, G. (2026). The development and validation of a critical thinking disposition scale for high school students. Acta Psychologica. https://doi.org/10.1016/j.actpsy.2025.106055

Luckita, I. A., Dewi, P., & Triana, J. (2025). English teachers's readiness in implementing deep learning approach at the elementary school level. **[Nama jurnal perlu diverifikasi]**, 12(2), 223–235. https://doi.org/10.37729/scripta.v12i2.6595

Riani, A., & Sujarwati, I. (2025). The preparedness of English teachers to implement deep learning in middle school. EDUCASIA: Jurnal Pendidikan, Pengajaran, dan Pembelajaran, 10(2), 229–244. https://doi.org/10.21462/educasia.v10i2.359

S. Tian, & Xiang, H. (2025). Exploring EFL students' effective use of GenAI in writing: Interplay of self-regulation, critical thinking, growth mindset, general writing competence, and GenAI literacy. Education and Information Technologies. https://doi.org/10.1007/s10639-025-13725-z

Saifulloh, A. I. (2021a). Ecocriticsm as the Development of Teaching Materials in Literary Theories and Appreciation Courses. *International Journal of English Education and Linguistics (IJoEEL)*, *3*(1), 1–9. https://doi.org/10.33650/ijoeel.v3i1.1752

Saifulloh, A. I. (2021b). *Kritik Sastra Lingkungan Asal Mula Teori, Pendekatan, dan Implementasi di Dunia Pendidikan*.

Saifulloh, A. I. (2026). *Introduction to Literature : Reading , Interpreting , and Analyzing Literary Texts*.

Saifulloh, A. I., & Anam, S. (2022). The Effectiveness of Pedagogical Ecocriticism to Enhance Environmental Awareness in Post Pandemic Era. *QALAMUNA: Jurnal Pendidikan, Sosial, Dan Agama*, *14*(1), 1–18. https://doi.org/10.37680/qalamuna.v14i1.1130

Saifulloh, A. I., Retnaningdyah, P., & Mustofa, A. (2025a). Immersive Situated Practice Within a Local Curriculum for Hybrid Learning. *Veredas Do Diretio*, 2–30. https://doi.org/https://doi.org/10.18623/rvd.v22.n5.3587

Saifulloh, A. I., Retnaningdyah, P., & Mustofa, A. (2025b). *Literacy as Resistance : Critical Framing for Promoting Ecological Awareness*. 996–1012. https://doi.org/10.19105/ghancaran.vi.21731

Saifulloh, A. I., Retnaningdyah, P., & Mustofa, A. (2025c). *Theoretical Praxis in Environmental Education within EFL Contexts*. *1*(1), 1–14.

Saifulloh, A. I., Retnaningdyah, P., Mustofa, A., & Suhartono, S. (2025). Immersive Situated Practice Within a Local Curriculum for Hybrid Learning. *Veredas Do Direito*, *22*(5), e223587. https://doi.org/10.18623/rvd.v22.n5.3587

Saifulloh, A. I., Retnaningdyah, P., & Mustopa, A. (2023). *Damaging Discourse in Online News Trends Over Ecological Literacy*. *5*(4), 885–900. https://www.elitejournal.org/index.php/ELITE

Saifulloh, A. I., Sengsouk, K., Fatoni, A., & Irfan, S. (2026). *Transforming Eco-Digital Literacies Through Multimodal Projects In English Language Learning*. *13*(01), 60–72.

W. Tian, Nasri, N. M., & Bin Jamaludin, K. A. (2026). An empirical study on the cultivation of critical thinking in high school academic English writing. https://doi.org/10.35516/hum.2026.8619

Walter, Y. (2024). Embracing the future of artificial intelligence in the classroom: The relevance of AI literacy, prompt engineering, and critical thinking in modern education. International Journal of Educational Technology in Higher Education. https://doi.org/10.1186/s41239-024-00448-3

Zhang, Y., Qi, W., Xia, C., Sun, H., & Chen, L. (2025). Exploring the effect of cooperative learning on senior high school students' critical thinking in EFL writing: An intervention study. Thinking Skills and Creativity. https://doi.org/10.1016/j.tsc.2025.101765

Zou, D., Xie, H., & Wang, F. L. (2022). Effects of technology enhanced peer, teacher and self-feedback on students' collaborative writing, critical thinking tendency and engagement in learning. Journal of Computing in Higher Education. https://doi.org/10.1007/s12528-022-09337-y