From Draft to Masterpiece: Using AI Writing Tools to Elevate form 6 ESL Essay Writing

Sathia Devi Thulasy., Dr. Nur Ainil Sulaiman

Faculty of Education, National University of Malaysia

DOI: https://doi.org/10.47772/IJRISS.2026.10100280

Received: 26 December 2025; Accepted: 01 January 2026; Published: 03 February 2026

ABSTRACT

Many Malaysian Form 6 ESL students struggle to achieve the MUET Writing Paper standards, especially in linguistic correctness, lexical range, and grammatical control. Addressing this issue, the current study looked into the efficacy of incorporating artificial intelligence (AI) writing tools to improve students' writing skills. The study, which was based on the Technology Acceptance Model (TAM) and Constructivist learning theory, looked at how AI-supported writing instruction helps learners engage and develop knowledge. A quantitative quasi-experimental approach was used with 16 Lower Six ESL pupils at a Malaysian secondary school. Over the course of a six-week intervention, students conducted guided MUET-style writing activities with AI writing tools. Pre- and post-test writings were graded using an analytical rubric that corresponded to MUET descriptors for vocabulary, grammar, and mechanics. Performance differences were measured using descriptive and inferential techniques. The findings demonstrated considerable improvements in all writing components, with vocabulary showing the biggest increase, followed by grammar and mechanics. These findings show that AI systems may efficiently scaffold linguistic accuracy and lexical development using instant feedback and iterative revision. The study emphasizes the pedagogical potential of AI-assisted instruction in improving MUET-oriented writing performance and advocates for its incorporation into post-secondary ESL curriculum.

Keywords: AI Tools, Writing skills, content, MUET

INTRODUCTION

The development of English writing proficiency remains a cornerstone of the Malaysian education system, particularly for Form 6 students preparing for the Malaysian University English Test (MUET). As a high-stakes assessment, the MUET Writing Paper (Paper 4) serves as a critical determinant of university admission and future employability, requiring candidates to demonstrate clarity, accuracy, and rhetorical cohesion within academic contexts (Tamimi et al., 2023). In alignment with global linguistic trends, Malaysia has adopted the Common European Framework of Reference for Languages (CEFR), which mandates that pre-university students achieve a minimum B2 competency level. This standard necessitates the ability to construct well-structured arguments and synthesize complex information skills that are essential for 21st-century academic success (Putra, 2023).

Despite these rigorous standards, Malaysian ESL learners frequently encounter significant obstacles in mastering academic writing. Empirical evidence suggests that student performance is often constrained by limited lexical range, grammatical inaccuracies, and underdeveloped critical thinking skills (Akhtar et al., 2019; Irvin & Blankenship, 2022). Furthermore, traditional pedagogical approaches often struggle to meet the needs of "digital-native" learners and fail to provide the immediate, individualized feedback necessary for iterative writing improvement. The demands of large-scale classroom management often prevent educators from delivering the tailored scaffolding required to address specific student deficits in coherence and organizational control (Yang et al., 2021).

To address these systemic challenges, national policies whereby including the *English Language Education Reform: The Roadmap 2015–2025* and the *Malaysia Digital Economy Blueprint (MyDIGITAL)* has advocated for the strategic integration of Artificial Intelligence (AI) to modernize language instruction. AI-powered tools

such as Grammarly and QuillBot have emerged as transformative resources, offering real-time grammatical correction, vocabulary expansion, and paraphrasing support through Natural Language Processing (NLP). Research indicates that these tools can enhance lexical variety and syntactic complexity while fostering learner autonomy and reducing the psychological barriers to writing (M. Nelfi & Norman, 2024; Rad et al., 2023).

However, a significant gap remains between the availability of these technologies and their practical application within the Malaysian Form 6 classroom. While international research highlights the benefits of AI-assisted instruction, there is a dearth of empirical data specifically focusing on its efficacy within the MUET framework. Many educators remain uncertain regarding the pedagogical integration of AI, risking a missed opportunity to align classroom practice with national digital goals. Consequently, this study seeks to investigate the impact of AI tools on the writing performance of Form 6 students in Pahang, exploring their perceptions and the challenges they face during the integration process. Specifically, the research addresses the extent to which these tools influence essay quality, the subjective experiences of the learners, and the practical hurdles encountered in an AI-augmented writing environment. To this end, the study aims the following research questions:

1. To what extent AI writing tools like Quillbot and Grammarly influence the writing performance of Form 6 ESL students in essay writing tasks?

2. What are the students' perceptions of using AI writing tools to enhance their essay writing?

3. What writing challenges do Form 6 ESL students face when integrating AI writing tools into their essay writing process?

Integrating AI tools like Grammarly and QuillBot into Form 6 ESL instruction offers a transformative solution to the logistical challenges of the high-stakes MUET, allowing educators to move beyond time-consuming manual marking toward a model of immediate, personalized scaffolding (Luo et al., 2025). By automating lower-order linguistic corrections, these digital systems enable instructors to prioritize higher-order cognitive development such as rhetorical strategy and critical thinking while simultaneously fostering learner autonomy among digital natives (Marza-Quispe et al., 2024; Paethrangsi et al., 2024). This pedagogical shift directly supports national strategic goals like *MyDIGITAL* and the *English Language Education Reform Roadmap*, providing a localized blueprint for operationalizing CEFR standards through technological innovation. Ultimately, the study bridges the digital divide in rural and semi-urban Malaysian settings, establishing a practical framework that ensures pre-university writing instruction meets the rigorous demands of 21st-century global academic literacy (Marzuki et al., 2023).

LITERATURE REVIEW

Sociocultural Theory in ESL Writing

Sociocultural Theory (SCT), rooted in Vygotskian principles, posits that language acquisition is a mediated process driven by social interaction, cultural artifacts, and scaffolding within the Zone of Proximal Development (ZPD). In the context of ESL instruction, SCT emphasizes that learners can transcend their individual linguistic limitations when supported by more capable peers or specialized resources, leading to documented improvements in cohesion, task response, and accuracy (Hong & D. B., 2020). When applied to modern pedagogy, AI writing tools function as sophisticated mediational means, serving as both scaffolding objects and interactive partners that provide real-time, modeled feedback. By viewing AI integration through the lens of SCT, this study evaluates how such digital tools facilitate cognitive and social development during the writing process, effectively bridging the gap between a student's current proficiency and their potential academic performance.

Cognitive Load Theory and AI Assistance

Cognitive Load Theory (CLT) posits that effective instructional design must manage the limitations of working memory by minimizing extraneous load to facilitate schema acquisition and germane processing. In the context of ESL writing, AI-driven technologies such as Grammarly and QuillBot where they can significantly reduce

the extraneous cognitive burden by automating lower-level tasks like grammar correction and lexical selection, thereby allowing learners to allocate more mental resources to higher-order rhetorical concerns, such as argumentative structure and audience awareness (Zhou et al., 2025).

However, recent research also cautions that unstructured or overly complex AI outputs may inadvertently increase superfluous processing if students are overwhelmed by the task of evaluating the suggestions. Consequently, a primary objective of this study is to determine whether AI integration effectively streamlines the writing process or if it introduces new cognitive demands that hinder the drafting and revision phases.

Technology Acceptance Model (TAM)

The Technology Acceptance Model (TAM) provides a robust framework for understanding how perceived usefulness and perceived ease of use determine a learner's behavioral intention and actual adoption of new technologies. Recent research extending TAM to generative AI and educational settings suggests that long-term integration is further influenced by factors such as pedagogy-fit, trust in the AI's output, and available institutional support (Mustofa & Rochman, 2025). Within the context of Form 6 ESL students, applying TAM allows for a deeper analysis of how personal attitudes, subjective norms such as teacher expectations and enabling conditions shape the decision to use AI writing tools. Ultimately, evaluating these perceptions of utility and usability is essential for explaining the variances in how students engage with AI and for determining the perceived learning value these tools provide during the writing process.

Integration of Theories in the Present Study

This study's conceptual framework takes an integrative approach, drawing on Sociocultural Theory (SCT), Cognitive Load Theory (CLT), and the Technology Acceptance Model (TAM), to investigate the pedagogical influence of AI writing tools on Form 6 ESL learners. Within this concept, AI writing tools serve as the independent variable, with instructor scaffolding acting as a mediator to facilitate learning within learners' zones of proximal development, as proposed by SCT. These technologies enable meaningful involvement with writing assignments by providing guided interaction and feedback.

Cognitive Load Theory explains how AI-assisted feedback and automated language support reduce extraneous cognitive load, allowing learners to devote more cognitive resources to germane processing tasks such as idea development, lexical selection, and syntactic accuracy. This cognitive optimisation is expected to result in improvements in the dependent variables, namely essay quality and linguistic accuracy in accordance with MUET writing descriptions.

In addition to these pedagogical and cognitive characteristics, TAM accounts for learners' acceptance and persistent use of AI writing tools by investigating perceived usefulness and perceived ease of use, which influence students' behavioural intention to use the technology. The paradigm also takes into account moderating variables such as prior linguistic skill, instructional tactics, and the MUET exam's high stakes. Overall, this integrated theoretical framework guarantees that the study remains pedagogically grounded while addressing cognitive efficiency and learner acceptance in the context of modern ESL writing instruction in Malaysia.

Writing Challenges Among Form 6 ESL Students

Pre-university Form 6 students must meet a number of challenging requirements, including discipline-specific genres, academic register, coherence across lengthier essays, and task completion for MUETs. Even at upper-secondary levels, empirical research continuously reveals enduring problems with lexical resources, cohesion/coherence, organisation, and argument development; these challenges are frequently made worse by a lack of metacognitive revision techniques and a lack of faith in academic registers. (Fitzgerald, A et al, 2025) These challenges were also evident among the Form 6 students who served as the sample for this study. Other frequently mentioned obstacles include a lack of clear training in revision techniques, feedback that only addresses superficial mistakes, and a significant mental strain during the planning and drafting stages. (Tutunaru, T 2023). Studies conducted nationally and regionally, especially in Malaysia, confirm that although students can achieve sufficient grammatical control, writing lengthy, persuasive essays is still difficult. These foundational

challenges influence the potential applications of AI tools, such as surface-level correction, rhetorical scaffolding, or external memory for reiterative drafts. (Tutunaru, T 2023).

Generative AI (e.g., QuillBot, ChatGPT, advanced paraphrasers) in ESL

Generative AI systems are able to create prompts, modify sections, and create drafts at the paragraph level. Recent empirical research and systematic reviews have started to map both the risks (academic integrity issues, decreased development of rhetorical planning skills if students accept outputs uncritically) and the promising effects (improved fluency, idea generation, draft quality when used as an ideation or revision aid). (Nobel, Lo et al, 2025) A recent comprehensive analysis of AI's effects on teaching academic writing paints a conflicting picture: Though careful pedagogical design is necessary to avoid superficial advances in merely surface qualities, AI can enhance instruction and reduce barriers to expression. (Sasson, I. et al 2021).

Gaps in the Literature

Despite the rapid proliferation of research on AI-assisted ESL writing, significant empirical gaps persist, particularly regarding the Malaysian pre-university context. Current literature predominantly focuses on either lower-secondary or tertiary students, leaving a notable void in evidence concerning Form 6 learners in government schools who must navigate advanced academic genres aligned with high-stakes MUET requirements (Mizan & Norman, 2024). Furthermore, most existing studies rely heavily on product-oriented metrics, such as final test scores, while neglecting process-oriented variables like planning time, the nature of revision moves, and the actual cognitive burden experienced by the learner (Putra, 2023). This imbalance makes it difficult to discern whether improvements in writing quality stem from genuine skill acquisition or a mere dependency on automated corrections.

Additionally, there is a lack of research addressing the intersection of pedagogical scaffolding and the long-term durability of AI-driven gains. Many studies fail to account for how local educational ecosystems which is shaped by the National Digital Education Policy and the CEFR-aligned English Roadmap that influence the actual adoption and sustainability of these tools (Khalifa & Albadawy, 2024). The current study seeks to address these deficiencies by employing a comprehensive mixed-methods approach that integrates process and product metrics with the Technology Acceptance Model (TAM). By investigating revision habits, cognitive load, and learner attitudes within a specific cohort in Pahang, the research aims to provide locally relevant insights into how AI can be strategically integrated to ensure meaningful skill transfer for students transitioning to higher education.

METHODOLOGY

This study employs a quasi-experimental pre-test and post-test design to evaluate the efficacy of AI writing tools in enhancing the essay writing skills of Form 6 ESL students. This methodology is particularly suited for educational research as it allows for the assessment of instructional interventions within authentic classroom settings where random assignment is often impractical (Creswell & Creswell, 2020). The research was conducted over a six-week period with a cohort of 16 students in Bentong, Pahang, involving an initial baseline assessment followed by an intervention phase where QuillBot and Grammarly were integrated into MUET Writing lessons. By comparing pre-test and post-test performance, the study captures quantifiable shifts in writing quality, aligning with contemporary trends in ESL research that utilize before-and-after comparisons to measure the impact of digital scaffolding on linguistic development (Mohamed K.MA, 2024).

Participants

The study focuses on a purposive sample of 16 Lower Six students from a government secondary school in Bentong, Pahang, a semi-urban setting representative of the diverse English proficiency levels found in the Malaysian pre-university system. These students are at a critical academic juncture, requiring CEFR B2/C1 competency to master the logical, analytical, and well-structured writing demanded by the high-stakes Malaysian University English Test (MUET). Despite these requirements, many learners struggle with organizational coherence and limited vocabulary, necessitating targeted intervention (Rosyada & A, 2023). The selection of a smaller, focused sample size is methodologically justified for classroom-based quasi-experimental research, as

it facilitates intensive monitoring of individual progress and detailed observation of how students engage with digital scaffolding (Vidhiasi, 2018). Equipped with necessary computer and internet access, this authentic classroom environment allows for a manageable yet in-depth exploration of how AI integration can bridge the gap between students' current drafting struggles and the rhetorical proficiency required for university-level success.

Instrument

The study utilized a dual-instrument approach to evaluate both writing performance and learner perceptions, centered on MUET Task 2 specifications. A quasi-experimental methodology was used to assess writing proficiency before and after the tests. Students wrote an unassisted argumentative essay as part of the pre-test to determine baseline proficiency in grammar, vocabulary, and organisation. Following a six-week intervention that included AI writing tools such as QuillBot and Grammarly, students completed a similar post-test essay challenge. Both writing samples were assessed using a standardised analytical rubric matched with the MUET Band Descriptors and CEFR Writing Scale to ensure construct validity and examination requirements. (Li et al., 2025). To complement these performance metrics, a structured questionnaire modified from validated AI literacy surveys (Biagini, 2024) was administered to capture student perspectives on the utility, ease of use, and impact of AI on their writing confidence. both the academic product and the affective dimensions of the writing process (Vivek et al., 2023). All quantitative data were analysed with the Statistical Package for the Social Sciences (SPSS) version 26. Descriptive statistics, such as means and standard deviations, were used to compare students' writing performance before and after the intervention. A paired-samples t-test was used to see if the observed changes were statistically significant for both total writing scores and specific components (vocabulary, grammar, and mechanics). Statistical significance was set at $p < .05$ level. Furthermore, Cohen's d was calculated to assess the magnitude of the intervention effect, allowing for an interpretation of practical significance beyond p-values. The combination of inferential statistics and effect size analysis allowed for a thorough assessment of the efficacy of AI-assisted writing education in improving MUET-oriented writing skills.

Data collection and Data Analysis

Data collection and analysis followed a structured six-week quantitative framework, beginning with a baseline pre-test in Week 1 to establish initial proficiency under exam conditions. During the subsequent four-week intervention phase, students engaged in guided lessons and independent tasks using QuillBot and Grammarly for brainstorming, paraphrasing, and grammatical refinement, culminating in a Week 6 post-test and an evaluative questionnaire. To measure the efficacy of this digital intervention, pre-test and post-test scores were analyzed using descriptive and inferential statistics, including mean and standard deviation comparisons, to identify significant shifts in grammar accuracy, vocabulary, and organization (Li et al., 2021). Parallel to these performance metrics, survey data were processed through frequency distributions, and mean scores to quantify student perceptions regarding the tools' utility and impact on writing confidence. This purely quantitative approach ensures that the findings are grounded in measurable trends, providing an objective evaluation of AI-assisted instruction within a high-stakes Malaysian educational context.

Findings

Descriptive Statistics of Writing Performance (RQ1)

Table 4.1 Descriptive Statistics for Pre-test and Post-test Writing Scores (n = 16)

Writing Components	Test	Minimum	Maximum	Mean	Standard Deviation
Task Fulfilment	Pretest	2	3	2.44	0.51
	Post test	3	6	4.56	0.81
	Pretest	1	3	2.19	0.66

Organisation & Coherence	Post test	3	6	4.31	0.79
Vocabulary	Pre test	1	3	1.81	0.54
	Post test	3	5	4.19	0.54
Grammar	Pretest	1	2	1.81	0.40
	Post test	3	5	4.06	0.68
Mechanics	Pretest	1	2	1.75	0.45
	Post test	3	5	3.75	0.68

Analysis of Writing Performance Improvements

According to Table 4.1, the descriptive analysis shows a significant rise in post-test mean scores across all five writing components when compared to baseline pre-test data. During the pre-test phase, the subjects had the lowest mean scores in Vocabulary (1.81), Grammar (1.81), and Mechanics (1.75), indicating severe baseline deficiencies in linguistic accuracy and adherence to writing rules. In contrast, the post-test data show a clear increasing trend, with the biggest mean improvements seen in Task Fulfilment (4.56) and Organisation and Coherence (4.31). These higher post-tests mean values indicate that the intervention improved the overall quality of the students' written outputs.

Despite the overall improvement in mean scores, the post-test findings show an increase in standard deviations for the majority of the components. For example, the standard deviation for Task Fulfilment grew from 0.51 to 0.81, whereas Grammar increased from 0.40 to 0.68. This increase in standard deviation indicates higher variability in student performance following the intervention. While the cohort's general competency improved, the disparity in scores shows that the degree of individual progress varied, with some students benefiting more strongly from the teaching than others.

In terms of mechanics, pupils made constant growth, with average scores rising from 1.75 (SD = 0.45) to 3.75 (SD = 0.68). Although the gains in mechanics were slower, the improvement is consistent with the findings of Missi, Mariam & Jarmouni et al, (2024), who found that persistent exposure to AI-supported revision procedures leads to higher correctness in punctuation, spelling, and formatting over time. Collectively, these findings show that the intervention was extremely effective in improving students' overall language competence, particularly in vocabulary use and grammatical accuracy, while also promoting considerable improvements in writing mechanics. Such findings add to the growing body of research demonstrating the pedagogical utility of AI technologies in improving ESL writing ability at the postsecondary level.

Inferential Analysis: Paired-Samples *t*-Tests and Effect Sizes

To evaluate if the observed differences between pre-test and post-test scores were statistically significant, paired-samples t-tests were performed for each writing component. Table 4.2 shows the complete inferential statistics, including t-values, degrees of freedom (df = 15), p-values, and effect sizes (Cohen's d).

Table 4.2 Paired-Samples *t*-Test Results for Writing Components (n = 16)

Writing Component	**Pre-test Mean**	**Post-test Mean**	**T**	**p**	**Cohen's *d***
Task Fulfilment	2.44	4.56	10.54	< .001	2.64
Organisation & Coherence	2.19	4.31	8.30	< .001	2.07

Vocabulary	1.81	4.19	19.00	< .001	4.75
Grammar	1.81	4.06	13.17	< .001	3.29
Mechanics	1.75	3.75	12.65	< .001	3.16

The paired-samples t-tests show significant differences in pre-test and post-test scores for all five writing components ($p < .001$). Effect sizes varied from big to very large, indicating significant changes in observed writing performance in this population (Cohen, 1988). The greatest effect was seen for vocabulary, followed by grammar and mechanics.

These effect sizes are understood as performance-based effects, suggesting improved essay outputs rather than clear evidence of long-term writing development. To assess the intervention's success, the findings were analysed using five fundamental writing components. The descriptive results show a universal improvement in all categories, with significant growth from pre-test to post-test.

Trends in Performance

When comparing what increased and decreased, the results demonstrate that the mean values for all writing sub-skills increased uniformly. Conversely, the frequency of low-level scores decreased; for example, while pre-test minimum scores for Vocabulary and Grammar were as low as 1, post-test minimums increased to 3. This trend indicates that the intervention effectively raised the "floor" of student performance, guaranteeing that even the worst writers produced higher-quality work.

Inferential Statistics

Paired-samples t-tests revealed statistically significant differences ($p < .001$) across all five components. With 15 degrees of variation, the study yielded big to very large impact sizes, with vocabulary displaying the most significant change. These findings are regarded as performance-based effects, indicating considerable improvements in the quality of the students' written essays in the research setting.

Correlation Analysis Between Pre-test and Post-test Performance

A Pearson correlation analysis was used to investigate the association between students' initial writing performance and post-test outcomes. Pearson correlation was found adequate because all variables were continuous and used the same rating criteria (Pallant, 2020). The research found a somewhat positive correlation between pre-test and post-test scores ($r = .42$, $p = .106$). The connection was not statistically significant at the.05 level.

This data implies that students' initial writing performance did not significantly predict post-test results. In practice, this means that students with lower pre-test scores were not systematically restricted to poorer post-test performance, and some lower-performing students made significant improvements. At the same time, the non-significant association highlights the variation in how students interacted with and benefited from the intervention.

Students' Perceptions of Using AI Writing Tools (RQ2)

Table 4.3 (Questionnaire)

Statement (Q)	Strongly disagree (%)	Disagree (%)	Neutral (%)	Agree (%)	Strongly Agree (%)
1.AI writing tools improved my overall essay writing	0	0	25	43.8	31.2

2. AI tools helped me organise my ideas better	0	0	12.5	50.0	37.5
3. AI tools helped improve my vocabulary	0	0	12.5	62.5	18.8
4. AI tools helped reduce grammatical errors	0	0	18.8	50.0	31.2
5. AI tools helped improve sentence clarity	0	0	18.8	56.2	25
6. I understand how to use AI tools ethically	0	0	12.5	62.5	25
7. I would recommend AI writing tools to other students	0	0	12.5	25	62.5
8. I am satisfied with the assistance provided by AI tools	0	0	6.2	50.0	43.8

Students' impressions were examined using frequency and percentage distributions based on replies to a five-point Likert scale questionnaire. Percentages are used to illustrate general response trends. Overall, pupils had positive attitudes about AI writing tools. 75.0% of students (43.8% agree; 31.2% strongly agree) said AI technologies helped them write better essays, while 25.0% were ambivalent. In terms of order and coherence, 87.5% of students agreed or strongly agreed that AI tools helped them better organise their ideas. Similarly, 81.3% noticed an improvement in language use, whereas just 6.2% disagreed.

In terms of acceptability and satisfaction, 81.2% of students said they would promote AI writing tools to other students, and 93.8% were pleased with the assistance provided. Additionally, 87.5% said they knew how to use AI tools ethically and responsibly. These findings suggest that students saw AI tools as beneficial, particularly for linguistic accuracy and revision, which is consistent with earlier research on learner views of automated writing feedback (Ranalli, 2021)

Summary of Findings

This chapter presented the results of pre-test and post-test analyses, correlation analysis, and questionnaire findings. Statistically significant improvements in writing performance were reported across all evaluated components, with large effect sizes indicating considerable performance changes within the sample. However, the non-significant link between pre-test and post-test scores, as well as the increased post-test variability, indicate that the intervention had varying degrees of success for children. Student perception data show that AI writing tools were perceived as useful and encouraging, particularly for vocabulary, grammar, and organisation.

DISCUSSION OF FINDINGS

Writing Performance and AI-Assisted Support (RQ1)

The study's findings show that AI writing tools like QuillBot and Grammarly were linked to measurable gains in students' essay writing performance within the context and duration of the intervention. These increases were evident in all five assessed writing components, including goal fulfilment, organisation and coherence, vocabulary, grammar, and mechanics, implying that AI-supported writing activities were associated with improved written outputs.

Paired-samples t-tests showed significant differences in pre-test and post-test scores for each writing component ($p < .001$). The extent of these differences was further substantiated by large to very large effect sizes (Cohen's d ranging from 2.07 to 4.75), indicating significant changes in observed performance within this group (Cohen 1988).

Based on the Cognitive Process Theory of Writing (Flower & Hayes, 2004), these findings might be understood as AI tools that aid in the reviewing and rewriting stages of the writing process by minimising the cognitive

burden associated with surface-level language decisions. The most significant performance gains were reported in vocabulary, grammar, and mechanics components, which correspond closely to the types of automated feedback supplied by AI writing tools. This pattern suggests that QuillBot and Grammarly had a significant impact on language accuracy and surface-level textual revision, which is consistent with previous research indicating that automated feedback systems are most effective at addressing form-focused aspects of writing (Ranalli, 2021; Zhang & Hyland, 2023). Importantly, such gains represent advances in written outputs rather than definitive evidence of profound linguistic internalisation.

Improvements were also seen in task completion, organisation, and coherence, showing that students were better able to reply to essay prompts and structure their ideas after the test. However, the rise in post-test variability indicates that these higher-order writing skills were not universally internalised by students. From a constructivist learning approach (Vygotsky, 1978), this uneven pattern suggests that AI tools served as conditional scaffolds, with their success determined by how learners engaged with and interpreted feedback.

Correlation analysis revealed that pre-test performance had no significant relationship with post-test performance, implying that initial writing skill did not strongly predict post-intervention outcomes. This finding implies that AI writing tools may have given temporary scaffolding to help students with lower initial proficiency, rather than exacerbating existing proficiency differences. Such a pattern is consistent with constructivist notions that mediated support can help learners perform beyond their current level, but it does not guarantee long-term skill transfer without ongoing guidance. (Taber, K, 2018).

Overall, the data indicate that AI writing tools had a moderate-to-strong impact on writing performance in this study, particularly by improving short-term writing outputs and revision quality. However, as per the Technology acceptability Model (Davis, 1989), positive performance outcomes and tool acceptability should not be construed as proof of autonomous learning efficacy. Rather, the impact of AI technologies appears to be context-dependent and pedagogically mediated, working best when combined with instructional guidance rather than as solo solutions (Ranalli, 2021).

Students' Perceptions and Technology Acceptance (RQ2)

The questionnaire results show that students had generally positive yet measured attitudes about the employment of AI writing tools in their essay writing process. A significant proportion of students indicated that AI tools were useful in assisting with writing assignments, particularly in terms of organisation, vocabulary, and general writing support. However, these judgements represent students' acceptance and perceived usefulness of the tools, not actual evidence of writing skill acquisition.
According to the Technology Acceptance Model (TAM) (Davis, 1989), the high percentages of agreement (75.0%) and satisfaction (93.8%) indicate significant perceived utility, a crucial factor of technology adoption. Students' willingness to propose AI writing tools to their colleagues (81.2%) demonstrates good acceptance in the classroom environment.

However, the occurrence of neutral replies across numerous questionnaire categories implies that not all students derived the same benefit or confidence from using AI tools. TAM does not make the assumption that acceptance always leads to learning benefits, and the data reflect this. While students judged AI tools to be helpful, these impressions do not demonstrate consistent increase in independent writing skills.
From a cognitive writing standpoint (Flower & Hayes, 1981), students' positive opinions may be linked to the role of AI tools in supporting the reviewing and rewriting stages of writing, which can make writing jobs more doable. Nonetheless, such facilitation may improve writing efficiency and confidence more than profound cognitive reorganisation of writing skills.

Overall, the findings show that students saw AI writing tools as valuable and supportive tools for essay writing, particularly for revision and language assignments. These impressions serve to explain students' involvement with AI tools, but they should be taken as markers of acceptance and usefulness, rather than direct evidence of learning success.

Challenges and Uneven Skill Internalization (RQ3)

Despite generally good perceptions and better writing skill, the findings highlight various problems in incorporating AI writing tools into students' essay writing processes. These issues are mostly related to uneven skill internalization, reliance on automated feedback, and variability in learning outcomes. Effective learning, according to constructivist learning theory (Vygotsky, 1978), necessitates active knowledge construction through mediated support. In this study, higher variability in post-test results implies that students engaged with AI-generated feedback in different ways. While some students appeared to use AI ideas as scaffolding to aid comprehension, others may have leaned too heavily on the tools without properly grasping the underlying writing principles. (Lixiang Yan, et al, 2025)

The concentrate of performance increases in vocabulary, grammar, and mechanics emphasises this problem. These components are tightly related to automated feedback, whereas higher-order skills like task completion and organisation, while better, showed larger dispersion. This pattern shows that students may have found it more difficult to transition from AI-supported revisions to independent control of discourse-level writing, a restriction identified in prior studies on automated feedback use (Ranalli, 2021; Zhang & Hyland, 2023).

Furthermore, the presence of indifferent responses in the questionnaire suggests that some students are unsure about how much AI tools helped them study. From a cognitive process standpoint, this could indicate disparities in students' capacity to interpret and integrate criticism throughout the revision stage of writing (Flower & Hayes, 1981). Without clear guidance, students may follow AI recommendations mindlessly, reducing opportunities for reflection and long-term learning.

Overall, the challenges revealed in this study indicate that AI writing tools work best as conditional scaffolding rather than autonomous learning solutions. Students need planned training to interact critically with AI input, minimise over-reliance, and develop transferable writing abilities. These findings support the notion that AI tools can improve writing ability, but their pedagogical usefulness is contingent on how they are integrated into teaching and learning environments.

Pedagogical and Educational Implications

Implications at the National Level

At the national level, the findings indicate that AI writing tools can help with continuing efforts to improve digital literacy and academic English competence among pre-university students. However, AI technologies should be positioned inside policy frameworks as supplemental educational supports rather than alternatives for teacher-led instruction.

Curriculum planners and policymakers should highlight AI literacy by explicitly teaching pupils how to evaluate, adjust, and reject AI-generated feedback. This is consistent with national digital education goals that emphasise responsible and critical use of technology over passive consumption (Davis, 1989)

Implications for Society and the Education System

From a societal standpoint, the findings emphasise the necessity of preparing students to engage critically with AI technologies, which are increasingly prevalent in academic and professional settings. While AI technologies might improve productivity and efficiency, the unequal internalisation shown in this study highlights the potential of over-dependence if utilised without proper direction. Educational institutions must consequently foster crucial digital abilities, ensuring that students grasp both the benefits and limitations of AI-assisted writing. This strategy promotes broader societal objectives such as ethical technology use and informed engagement in a digital economy.

Implications for Teachers and Students

For teachers, the data indicate that AI writing tools are most effective when used as guided scaffolding inside organised writing instruction. Teachers should mix AI-assisted draughting with opportunities for AI-independent

writing, reflection, and feedback to foster deeper learning and skill transfer. AI tools can help pupils improve their revising skills and linguistic accuracy. However, rather than blindly accepting AI criticism, students should be pushed to engage critically with it. Developing this evaluative skill is critical for promoting independent writing competence (Flower & Hayes, 1981; Vygotsky, 1978).

Limitations of the Study

Several restrictions should be noted. First, the study had a limited sample size (n = 16), limiting the generalisability of the results. Second, the intervention was relatively brief, therefore the results represent short-term performance outcomes rather than long-term writing growth. Third, the study was mostly based on pre-test, post-test, and questionnaire results. While these instruments gave useful insights into performance and perception, they could not capture the in-depth cognitive processes that underpin writing development. Future research could use qualitative methodologies to investigate how students interpret and use AI feedback over time.

CONCLUSION

This study looked at the impact of AI writing tools on the essay writing performance of Form 6 ESL students. The findings show that AI tools like QuillBot and Grammarly were related with statistically significant gains in writing skills and were generally well received by students. However, the study also shows issues such as uneven skill internalisation and reliance on automated feedback. Based on Cognitive Process Theory, Constructivist Learning Theory, and the Technology Acceptance Model, the study views AI writing tools as conditional and context-dependent scaffolding. While they can improve short-term writing productivity and revision, they cannot ensure long-term writing skill improvement. To summarise, AI writing tools have pedagogical potential when strategically integrated into instructional frameworks that value critical engagement, instructor mediation, and opportunities for independent writing. Their value is not in replacing traditional education, but in complementing it in ways that improve both performance and learning in ESL classrooms.

REFERENCES

1. Akhtar, R., Hassan, H., Saidalvi, A. & Hussain, S. (2019). A systematic review of the challenges and solutions of ESL students' academic writing. International Journal of Engineering and Advanced Technology, 8(5C).
2. Biagini, G., Cuomo, S. & Ranieri, M. (2024). Developing and validating a multidimensional AI literacy questionnaire. ((PDF) Developing and Validating a Multidimensional AI Literacy Questionnaire: Operationalizing AI Literacy for Higher Education)
3. Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
4. Creswell, J. W., & Creswell, J. D. (2020). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches (5th ed.). SAGE.
5. Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and technology acceptance. MIS Quarterly, 13(3), 319–340. DOI:10.2307/249008
6. Fitzgerald, A., Avirmed, T., & Battulga, N. (2025). Exploring the factors informing educational inequality in higher education: a systematic literature review. Perspectives: Policy and Practice in Higher Education, 29(4), 199–209. https://doi.org/10.1080/13603108.2024.2381121
7. Flower, Linda & Hayes, John. (2004). A Cognitive Process Theory of Writing. College Composition and Communication. 32. DOI:10.2307/356600
8. Hong, N., & del Busto, C. T. (2020). Collaboration, scaffolding, and successive approximations: A developmental science approach to training in clinical psychology. Training and Education in Professional Psychology, 14(3), 228–234. https://doi.org/10.1037/tep0000260
9. Irvin, Sandra and Blankenship, Kathryn Guillot (2022) "Vocabulary & Academic Success in University Undergraduate Students," Teaching and Learning in Communication Sciences & Disorders: Vol. 6: Iss. 2, Article 6. Available at: https://ir.library.illinoisstate.edu/tlcsd/vol6/iss2/6

10. Khalifa, Mohamed & Albadawy, Mona. (2024). Using Artificial Intelligence in Academic Writing and Research: An Essential Productivity Tool. Computer Methods and Programs in Biomedicine Update. DOI:10.1016/j.cmpbup.2024.100145
11. Lihua Li, Meaghan S Cuerden, Bian Liu, Salimah Shariff, Arsh K Jain & Madhu Mazumdar (2021) Three Statistical Approaches for Assessment of Intervention Effects: A Primer for Practitioners, Risk Management and Healthcare Policy, , 757-770, DOI: 10.2147/ RMHP.S27583. https://doi.org/10.2147/RMHP.S275831
12. Li, C., Nik Mohd Alwi, N. A. & Azmat Ali, M. M. (2025). A comparative review of CET4 essay assessments. Asian Journal of English Language and Pedagogy, 13(1), 101–118. (https://umpir.ump.edu.my/id/eprint/45169/1/A%20comparative%20review%20of%20CET4%20essay%20writing%20assessments%20insights%20from%20the%20CEFR%20level%20descriptors%20and%20CET4%20writing%20rubrics.pdf)
13. Lixiang Yan, Roberto Martinez-Maldonado, Yueqiao Jin, Vanessa Echeverria, Mikaela Milesi et al, (2025)The effects of generative AI agents and scaffolding on enhancing students' comprehension of visual learning analytics,Computers & Education, Volume 234, https://doi.org/10.1016/j.compedu.2025.105322.
14. Luo, L., Hu, J., Zheng, Y. et al. (2025). Human vs. AI: Does AI learning assistant enhance students' innovation behavior?. Educ Inf Technol 30, 17483–17530 https://doi.org/10.1007/s10639-025-13474-z
15. Malaysia Digital Economy Blueprint (https://naqiz.com/wp-content/uploads/2021/02/MyDigital-Malaysia_Digital_Economy_Blueprint.pdf)
16. Maraza-Quispe, B. et al. (2024). Impact of gamified online tools. International Journal of Information and Education Technology, 14(2), 123–130. Doi: 10.18178/ijiet.2024.14.1.2033
17. Marzuki, , Widiati, U., Rusdin, D., Darwin, , & Indrawati, I. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. Cogent Education, 10(2). https://doi.org/10.1080/2331186X.2023.2236469
18. Missi, Mariam & Jarmouni, Rania & Chibi, Mounir & Azize, El. (2024). Exploring Students' Perceptions of Using Grammarly in Academic Writing: Master Students of the English Department as a Case Study. International Journal of Innovative Research and Development. 13. 24-31. DOI:10.24940/ijird/2024/v13/i6/JUN24003
19. Mizan, Nelfi & Norman, Helmi. (2024). Pre-University Students' Perception in Using Generative AI: A Study at a Malaysian Private University. International Journal of Academic Research in Business and Social Sciences. DOI:10.6007/IJARBSS/v14-i8/22455
20. Mohamed K, M A (2024) Using artificial intelligence in academic writing and research: An essential productivity tool, Computer Methods and Programs in Biomedicine Update, DOI:10.1016/j.cmpbup.2024.100145
21. Mohd Don, Zuraidah. (2015). English Language Education Reform in Malaysia: The Roadmap 2015-2025.(https://www.researchgate.net/publication/290260030_English_Language_Education_Reform_in_Malaysia_The_Roadmap_2015-2025)
22. Mustofa, R. H. et al. (2025). Extending TAM: Role of trust and ethics. Computers and Education: AI, 8. https://doi.org/10.1016/j.caeai.2025.100379.
23. Noble Lo, Alan Wong, Sumie Chan (2025) The impact of generative AI on essay revisions and student engagement, Computers and Education Open, Volume 9, https://doi.org/10.1016/j.caeo.2025.100249.
24. Paethrangsi, Nisara & Teekasap, Sombat & Khiewpan, Rachapong & Jandaboue, Wanlapa. (2024). Empowering Students' Autonomous Learning through Self-regulation, Metacognitive Strategies, and Collaborative Learning Environments. Journal of Liberal Arts RMUTT. 5. 69-79. DOI:10.60101/jla.2024.5.1.4065
25. Pallant, J. (2020). SPSS Survival Manual: A Step by Step Guide to Data Analysis Using IBM SPSS. London: McGraw-Hill, Open University Press. https://doi.org/10.4324/9781003117452
26. Putra, M. (2023). Teachers' and students' perceptions on AI correction tools. Jurnal Tatsqif, 21, 35–66. DOI:10.20414/jtq.v21i1.7963
27. Rad, H. S., Alipour, R., & Jafarpour, A. (2024). Using artificial intelligence to foster students' writing feedback literacy, engagement, and outcome: A case of Wordtune application. Interactive Learning Environments, 32(9), 5020–5040. https://doi.org/10.1080/10494820.2023.2208170

28. Ranalli, Jim (2021) L2 student engagement with automated feedback on writing: Potential for learning and issues of trust, Journal of Second Language Writing, Volume 52, https://doi.org/10.1016/j.jslw.2021.100816.
29. Rosyada-AS, Amrina & Apoko, Tri. (2023). Investigating English Vocabulary Difficulties and Its Learning Strategies of Lower Secondary School Students. Journal of Languages and Language Teaching. 11. 489. 10.33394/jollt.v11i3.8404. DOI:10.33394/jollt.v11i3.8404
30. Sasson, Irit & Yehuda, Itamar & Miedijensky, Shirley & Malkinson, Noam. (2021). Designing new learning environments: An innovative pedagogical perspective. The Curriculum Journal. DOI:10.1002/curj.125
31. Taber, Keith. (2018). Scaffolding learning: Principles for effective teaching and the design of classroom resources. (https://www.researchgate.net/publication/327833000_Scaffolding_learning_Principles_for_effective_teaching_and_the_design_of_classroom_resources)
32. Tamimi, N. N. M., Al O.M.S., & Thottoli, M. (2023). Factors influencing university selection. Journal of Business Management Review, 4, 439–454. DOI:10.47153/jbmr46.7142023
33. Tutunaru, Tatiana. (2023). Improving Assessment and Feedback in the Learning Process: Directions and Best Practices. Research and Education. 38-60. 10.56177/red.8.2023.art.3. (https://researchandeducation.ro/wp-content/uploads/2024/03/Tatiana_Tutunaru_Improving_Assessment_and_Feedback_in_the_Learning_Process_Directions_and_Best_Practices_REd_nr8.pdf)
34. Vidhiasi, Dhion. (2018). Classroom observation and Research. DOI:10.6084/m9.figshare.7988522
35. Vivek, R., Nanthagopan, Y. & Piriyatharshan, S. (2023). Theoretical underpinnings of triangulation. SEEU Review, 18.Volume 5, https://doi.org/10.1016/j.cmpbup.2024.100145.
36. Vygotsky, L. S. (1978). Mind in Society: Development of Higher Psychological Processes (M. Cole, V. Jolm-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press. (https://www.scribd.com/document/845099157/Vygotsky-1978-Mind-in-Society-Development-of-Higher-Psycholo)
37. Yang, L., Chiu, M. M. & Yan, Z. (2021). Power of teacher feedback. Educational Psychology, 41(7), 821–824. DOI:10.1080/01443410.2021.1964855
38. Zhang, Z., & Hyland, K. (2023). Student engagement with peer feedback in L2 writing: Insights from reflective journaling and revising practices. Assessing Writing, 58, Article 100784. https://doi.org/10.1016/j.asw.2023.100784
39. Zhou, X., Li, Y., Chai, C. S., & Chiu, T. K. F. (2025). Defining, enhancing, and assessing artificial intelligence literacy and competency in K-12 education from a systematic review. Interactive Learning Environments, 33(10), 5766–5788. https://doi.org/10.1080/10494820.2025.2487538