



International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: <https://www.svedbergopen.com/>



Research Paper

Open Access

AI-Assisted Metacognitive Disconnect a Study of Female EFL Learners' Writing

Enas Abdelwahab Eltom RahmtAllah

Department of English Language & Literature, College of Languages & Humanities, Qassim University, Unaizah P.O.Box 5380, Saudi Arabia

E.RAHAMAT@qu.edu.sa

<https://orcid.org/0009-0009-0725-7128>

Abstract

This study compares the relationship between EFL learners' metacognitive awareness of the strengths, weaknesses, and the quality of their writing under AI-assisted versus non-AI-assisted conditions. The study used a quasi-experimental, repeated-measures design with 63 university-level EFL learners who completed argumentative writing tasks under both conditions in a common timeframe. Two EFL professionals evaluated the anonymized submissions after each writing session using a rubric comprising four criteria. Results indicated significant benefits of artificial intelligence (AI) in essay writing with disparate scores between learners' performance when they utilized AI ($M = 11.06$) compared to when there was no AI support ($M = 8.21$), with an effect size of $d = 1.29$. Survey results indicated learners' satisfaction with the results achieved using AI ($M = 7.25$) compared to those achieved without it ($M = 5.48$). From the evaluators' perspective, however, there was no significant qualitative difference between the two sets of essays. Another significant finding pertained to learners' perception of AI as a tool that improved the quality of their writing though with concerns of over-dependence on it which can be inferred as an indication of metacognitive disconnect.

Keywords: Artificial Intelligence (AI); Metacognitive awareness; Metacognitive Disconnect; Self-Assessment; Writing Performance

1. INTRODUCTION

The evolution of generative AI tools such as ChatGPT and Grammarly has facilitated correction and editing of texts across academic subjects and levels. Given their many benefits, these tools are being adopted in the teaching of English as a foreign language, particularly in the development of writing. Apart from the secondary function of editing, these tools are being extensively used for producing content on a large scale [7,10, 20]. In the current educational landscape, the rapid development and widespread adoption of artificial intelligence tools that function as metacognitive mirrors offer new opportunities for supporting learners' metacognitive processes. They encourage creativity, innovation, and reasoning in learners, thereby contributing to the development of writing processes such as, reviewing, evaluating, and correcting errors [24,25, 26]. However, these tools represent a double-edged sword. Although they can empower learners to produce writing that is accurate in terms of grammar, style, and narrative structure, they may also generate psychological and metacognitive pressures [5, 7, 10].

Metacognitive disconnection has been identified as a significant factor influencing learners' self-efficacy and motivation, acting as an obstacle that affects writing quality and the ability to improve it [7]. Several recent studies have addressed and discussed specific aspects of this phenomenon highlighting the 'Regulatory Gap' between metacognitive awareness and its actual enactment during writing [26], Critical-AI Gap between the technical use of AI tools and the fulfillment of a task's rhetorical purpose [26], and Affective-Performance Discrepancy, whereby reported motivational gains do not correlate with measurable performance outcomes [7]. Psychological and affective factors influence learners' motivation when making spontaneous educational decisions in the EFL classrooms. These factors create time-constrained conditions during tests, prompting learners to review their written work in a hurried and unsystematic manner, which may hinder methodical revision. Furthermore, affective

states can intensify this tendency to rush, potentially leading learners to adopt less deliberate revision strategies [14, 22]. Few studies have examined how time-constrained metacognitive processes operate in the context of AI-assisted English language learning in an EFL classroom, a gap which the present study aims to bridge. The study, hence, addresses the following research questions:

1. In the context of EFL, is there a relationship between learners' strengths and weaknesses during timed essay writing and metacognitive disconnect?
2. How does the quality of writing compare in the presence versus absence of AI tools?

2. THEORETICAL FRAMEWORK

2.1. Metacognition and the Foundations of Metacognitive Disconnect

The three processes involved in writing—planning, monitoring, and evaluating—serve as the foundational for a tripartite model of metacognition [28]. Each process addresses a specific aspect: planning, for instance, involves organizing ideas and setting goals. Monitoring entails tracking progress, identifying errors, and addressing them. Finally, evaluation determines the quality and effectiveness of the final product, while also considering the methods employed to achieve it. This tripartite model provides a basis for examining the extent to which learners may accurately identify the strengths and weaknesses in their own writing.

Based on this foundation, several established concepts describe situations where these subjective judgments fail to track actual performance. Metacognitive inaccuracy refers to the failure to accurately assess one's own judgments or performance. The failure may result from inadequate self-assessment skills, motivational biases, difficulty in evaluating previous performance [1, 30], or the Dunning-Kruger effect which is a tendency for individuals with lower levels of ability to overestimate their competence, whereas those with higher levels of ability may underestimate theirs [15, 18].

Similarly, the concept of the illusion of knowing describes situations in which learners are overconfident about their levels of understanding. Such judgments may be influenced by task characteristics, including text length and question format [3,15]. Further, the gap between confidence during study and actual test performance may arise because learners rely on cues that are available during study but absent during retrieval [2, 6, 17]. [8] provided further support to this by highlighting significantly the gaps between reported humility and actual perceptual accuracy in tasks. To sum up, cognitive psychology highlighted the disconnect between perceived and actual performance as a long-recognized phenomenon not requiring a new label. Thus, it is the body of work that the present study draws on in defining metacognitive disconnect as the discrepancy between learners' beliefs about what they know or can do and their objectively measured performance.

2.2. AI-Assisted Metacognitive Disconnect: Extending the Construct

A recent line of research directly addresses this divergence in human-AI interaction. [16] proposed an AI-induced metacognitive breakdown, describing a widening gap across four variables—outcome, perception, calibration, and competence—as reliance on AI increases. [13] documented performance-belief gaps of up to 34.6 percent between self-confidence and objective accuracy. They also identify "confirmation valuation," a phenomenon where reliance on AI during highly difficult tasks increases even as confidence in the output and objective performance decline. [12] described a related metacognitive tension between using AI to facilitate thinking versus substituting for it, while [19] framed the situation in terms of calibration and misalignment, arguing that generative AI can improve work outcomes while simultaneously reducing the accuracy of self-assessment. Similarly, the study by [23] involving novice programmers reported an "illusion of competence," where struggling learners perceive high performance despite heavy reliance on AI-generated output.

Overall, these studies characterize the phenomenon as a form of cognitive offloading, wherein users delegate cognitive tasks to external systems. They highlight counterintuitive findings suggesting that high levels of AI literacy are sometimes associated with less accurate self-assessment, as tech-savvy users may become overconfident when evaluating AI-assisted outcomes [11]. Consequently, this research defines AI-assisted metacognitive misalignment as the discrepancy between a learner's perceived understanding or ability and their actual knowledge or competence when working independently. This occurs when AI-generated output enhances immediate task performance or creates an illusion of understanding, without a corresponding improvement in the learner's fundamental comprehension, knowledge transfer, or ability to verify the result [11, 12,13,16,19, 23]. This recent debate on AI in education, where the need to preserve learner agency in the context of personalized and adaptive learning is highlighted. Excessive dependence on automated systems may create a gap between immediate performance support and the development of independent learning capabilities [31].

2.3. Application to EFL Writing Under Timed Conditions

In the field of EFL writing research, a general framework helps elucidate various gaps identified in recent studies. [26] identified a gap between students' metacognitive awareness and its actual application during writing, as well

as an "AI-criticality gap" between the technological use of AI tools and the achievement of the writing task's communicative goal. [7] also observed a discrepancy between engagement and performance, where high levels of student effort and motivation associated with AI use did not translate into tangible improvements in writing quality. At the same time, AI tools can serve as "metacognitive mirrors," enhancing students' understanding when the interaction is structured and reflective, thereby enabling them to make more accurate judgments over time [26, 29]. Likewise, recent studies on AI-assisted academic tutoring emphasize the capabilities of LLMs in offering interactive learning support, content creation, and knowledge enhancement. These applications highlight the need to view AI as a tool to enhance learning, not replace it, and to encourage students to engage with the material on their own [32].

It appears that time pressure exacerbates these changes. Under time constraints, students' revision behaviors may become more erratic and less systematic, driven partly by affective states such as anxiety rather than by deliberate revision strategies [14, 22]. Conversely, other studies have shown that AI-assisted feedback improves argumentative writing performance under time constraints compared to unassisted writing or direct teacher feedback. It is within this context of heightened time pressure and AI mediation that the present study examines EFL students' awareness of their writing strengths and weaknesses, specifically, in comparison with writing produced without AI assistance.

2.4. Previous Studies

[26] reported that AI-assisted reasoning can reveal a communication hurdle: students may use an AI tool appropriately yet fail to achieve the key linguistic objective of the task. This finding is closely aligned with the present research which distinguishes between proficiency in using AI and genuine communicative understanding by examining the discrepancy between students' self-assessed writing quality and their objectively assessed performance.

[7] tracked low-proficiency EFL students using an AI-powered chatbot and found that although affective benefits such as motivation and engagement increased following AI use, actual writing performance improved primarily in terms of linguistic complexity. Their concept of a "readiness gap"—whereby learners' proficiency must reach a certain threshold before they can effectively utilize AI capabilities—offers a possible explanation for the individual differences in cognitive processing observed in the present study.

[13] conducted a three-stage study of human problem-solving assisted by AI and found that trust and reliance on AI increased even when actual performance and verification accuracy declined during highly challenging tasks. Their proposed 'verification paradox' provides a possible mechanism for the widening of the metacognitive gap, particularly when tasks are highly challenging, as was the case under the time-constrained conditions examined in the present study.

[16] re-examined the Dunning-Kruger analogy in the context of AI assistance and proposed that AI can decouple four variables that typically co-occur in unaided performance namely, product quality, self-perception, calibration, and latent ability. This interpretation supports the view that AI-mediated calibration may constitute a distinct dimension of metacognitive judgment rather than simply reproducing conventional forms of human miscalibration.

[29, 34, 35] investigated the impact of an Automated Writing Evaluation (AWE) tool on metacognitive engagement in persuasive writing. They reported that sustained interaction with AWE feedback prompted students to recalibrate their perceptions over time, leading to heightened awareness of text structure and grammar. These findings are consistent with the present framework's proposition that AI feedback can function as a 'metacognitive mirror,' potentially narrowing the gap between self-assessment and performance as interactions become sustained and reflective.

[14] integrated a theoretical framework of metacognition—addressing metacognitive knowledge, experience, and strategies across the planning, monitoring, and evaluation stages of EFL writing—by surveying 82 Croatian high school students to assess their strategy use and attitudes toward ChatGPT. The results indicated that students demonstrated strong metacognitive awareness during the planning stage but were less consistent in employing strategies during the post-writing phase. Furthermore, the use of ChatGPT was largely confined to the pre-writing stage, particularly for brainstorming and drafting, where students held positive, albeit cautious, attitudes toward the tool. This discrepancy between strong awareness during the planning and less consistent strategy use during monitoring and evaluation closely parallels the regulation gap described in the theoretical framework. It suggests that within the context examined in this study difficulties in metacognitive regulation may become more apparent during the later, revision-oriented stages of writing than during initial planning.

[9] piloted a four-stage, AI-integrated process-writing intervention consisting of self-assessment via rubric, AI-assisted evaluation, AI-informed revision, and reflection with 35 university EFL pedagogy students at B1 and C1

proficiency levels. Combining a structured feedback form with inductive thematic analysis of student reflections, the study found that the participants reported benefits including immediate feedback, heightened metacognitive awareness, and perceived improvement in writing quality. B1 students particularly valued grammatical and structural support whereas C1 students valued assistance with argumentation and coherence. Because the study measured perceived awareness and perceived improvement, it illustrates the limitations of self-report evidence that the present study seeks to address by pairing self-assessment with objective performance measures.

[10] randomly assigned 60 Iranian intermediate EFL learners to receive either AI-mediated feedback, delivered through the Poe application, or traditional teacher feedback, and assessed writing proficiency using IELTS Writing Task 2 as pre- and post-tests. Students who received AI-generated feedback demonstrated significant improvements in grammatical accuracy and coherence compared with those who received teacher-based feedback. Follow-up discussions revealed that students valued the speed, personalization, and accessibility of AI feedback, reporting that it enhanced their motivation and autonomy. However, participants expressed concerns about the loss of human interaction, the risk of over-reliance, and the tool's limited sensitivity to contextual nuances. These quantitative benefits and qualitative reservations highlight the distinction—underscored by the current framework—between AI's tangible advantages in enhancing objective accuracy and the more complex processes of student self-regulation and metacognitive judgment.

[24] conducted a 15-week action research study involving 148 undergraduate students at Bahauddin Zakariya University in Pakistan. The experimental group used a suite of AI tools including ChatGPT, Claude AI, Meta AI, and Canva, within a structured “cognitive partnership” model, whereas the control group received traditional, teacher-led instruction. The experimental group achieved statistically significant gains in writing skills ($Z = -6.325$, $p < .001$), while the control group also registered comparatively modest but statistically significant gains ($Z = -2.128$, $p = .033$). Notably, substantial skill improvement was observed only after six to eight weeks of continuous use. Qualitative analysis further indicated that participants perceived AI tools as cognitive partners, mirrors for metacognitive processes, and instruments for fostering equity. The majority of participants (79.7%) reported positive perceptions, and 86% expressed a willingness to continue using these tools. The researchers concluded that an educational framework incorporating collaboration with AI was an important factor in achieving successful and ethical outcomes. This finding is consistent with the present framework's proposition that structured and reflective collaboration with AI may help reduce metacognitive discrepancies.

[28] conducted an eight-week mixed-methods study with 40 university students in eastern China. Participants were randomly assigned to two groups: one used a generative AI chatbot namely, the Tongyi.ai platform, while the control group used traditional resources. Outcomes were measured using the Metacognitive Awareness Inventory, the Strategy Inventory for Language Learning, writing performance, interaction logs, reflective journals, and interviews. The AI group demonstrated greater improvement than the control group in task strategy diversity, metacognitive awareness, and writing performance. Qualitative data indicated that the chatbot supported task planning, encouraged flexible and adaptive strategies, and enhanced metacognitive thinking, although participants also expressed concerns about over-reliance on the tool and the accuracy of its outputs. These findings provide empirical support for the three-component model of planning, monitoring, and evaluation that strengthens the present framework, while the participants' concerns about over-reliance also reflect the tension between performance gains and the potential risks of inaccurate metacognitive judgments.

[6] and [30], whose research predates the emergence of generative AI, established that illusions of competence and metacognitive inaccuracy can occur in self-regulated learning even in the absence of technological mediation. Their work provides an important baseline for the present study: it indicates that any metacognitive disconnect observed in the AI-assisted condition may reflect ordinary human miscalibration rather than AI use itself.

In conclusion, the literature presents a layered understanding of metacognitive disconnect, extending from well-established forms of human miscalibration to emerging AI-related mechanisms such as cognitive offloading, variable decoupling, and the verification paradox, to documented manifestations in EFL writing including regulatory and the affective-performance gaps. Taken together, these strands provide a rationale for examining metacognitive disconnect specifically in timed, AI-assisted writing, where cognitive, affective, and technological factors may interact to influence learners.

3. METHODOLOGY

3.1 Research Design

This study uses a quasi-experimental design to compare the quality of English writing of EFL learners under two distinct conditions namely, with and without AI assistance. A dialectical writing task which requires the presentation of arguments for and against a position was conducted in both conditions to isolate the impact of AI use on writing and facilitate a direct comparison of results. Surveys before and after the study helped capture the perceptions so that they could be correlated with actual performance.

3.2 Participants

Initially the number of participants in the study was ninety-seven, all female learners enrolled in the seventh and eighth levels of the Department of English Language and Literature at the College of Languages and Humanities, Unaizah, Qassim University. Out of these, thirty-two were excluded from the final sample. Fifteen of these students participated in a pilot study and were excluded from the study to avoid skewing the original sample. The remaining seventeen were excluded for failing to meet all study requirements. The failure refers to not completing questionnaires or tests, or both. Consequently, the final study sample consisted of sixty-five participants.

Before being enrolled at the English department, all participants had completed preparatory English language courses from level one to level six which included courses in communication skills and translation, creating a robust foundation for their language development.

3.3 Instruments

The study used three instruments for data collection: i) a controlled essay writing task, ii) a writing assessment rubric, and iii) questionnaires administered before and after the writing tasks.

3.3.1 Essay Writing Task

All participants completed an English writing task first without AI-assistance and thereafter with the freedom to use any generative AI tool of their choice. The exercises were to be completed within the stipulated timeframe of 45 minutes each.

3.3.2 Writing Assessment Rubric

Essays from both the exercises were evaluated for vocabulary usage, guiding instructions, information sources application, grammatical accuracy, and writing proficiency. Each dimension was scored on a scale (1 = poor, 5 = excellent) with a maximum possible total of twenty points.

3.3.3 Questionnaires

Two questionnaires were administered immediately after each writing task in the study to gather comprehensive data on participants' perceptions and experiences with the two writing modalities.

4. RESULTS AND DISCUSSION

4.1 Preliminary Analyses

The final group included in the statistical analysis consisted of 63 individuals with usable writing output (two were excluded during data processing because their essays failed to reach the minimum word limit).

4.2 Writing Quality: Instructor-Assigned Scores Under AI-Assisted versus Non-AI-Assisted Conditions

A paired-samples t-test compared the instructor-assigned scores across conditions. Essays written with AI assistance received scores higher ($M = 11.06$, $SD = 2.05$) than essays written without AI assistance ($M = 8.21$, $SD = 1.97$), $t(62) = -10.25$, $p < .001$, 95% CI for the mean difference $[-3.41, -2.30]$, Cohen's $d = 1.29$ (a large effect), confirmed by a Wilcoxon signed-rank test ($p < .001$).

These results indicate that AI assistance led to a visible and measurable improvement in writing quality of the EFL learners. This finding aligns with previous studies that concluded that AI-generated suggestions (even if imperfect) may increase linguistic accuracy and coherence (Elghani & Asadzadeh-Maleki [10]; Rashid et al. [25]).

4.3 The Metacognitive Gap: Self-Overestimation With and Without AI Assistance

The central test of the study's research question compares each participant's metacognitive gap — the extent to which her self-predicted score exceeded her actual, instructor-assigned score — between the two writing conditions. Under non-AI conditions, participants overestimated their performance by an average of 5.48 points ($SD = 2.9$). With AI support, these self-assessment ratings increased significantly, reaching a mean score of 7.25 ($SD = 2.37$ for the AI group). Comparative analysis confirmed the statistical significance of this difference— $t(62) = -4.50$, $p < .001$, 95% CI of the difference $[-2.57, -0.99]$, Cohen's $d = 0.57$ (medium effect size)—a result further supported by the Wilcoxon test ($p < 0.0$). In other words, although the AI assistance improved writing quality, participants perceived this benefit more than the actual improvement achieved, even surpassing the tendency to overestimate their unaided writing abilities. This provides clear statistical evidence of the difference between perception and reality when using AI: the benefits of perceived AI exceed its objectively measurable value.

4.4 Relationship Between the Non-AI and AI-Assisted Metacognitive Gaps

To determine whether the tendency to overestimate performance capabilities is an intrinsic trait or a context-dependent phenomenon, the conditions during the non-AI and the AI-supported writing were directly compared. The two gap scores were positively but only weakly-to-moderately associated, Pearson $r(61) = .314$, $p = .012$ (Spearman $\rho = .282$, $p = .025$). Participants who overestimated their non-AI performance more also tended to

overestimate their AI-assisted performance, but the association was modest, indicating that overestimation is only partly a stable trait; a substantial portion of each participant's degree of overestimation appears to be specific to the particular writing condition.

4.5 Calibration Accuracy Within Each Writing Condition

A complementary way to assess metacognitive accuracy is to correlate predicted and actual scores within each condition. Under non-AI conditions, predicted (PS) and actual (AS) scores were positively correlated, $r(61) = .338$, $p = .007$ ($\rho = .400$, $p = .001$). Under AI-assisted conditions, predicted (PS-AI) and actual (AS-AI) scores showed an almost identical correlation, $r(61) = .338$, $p = .007$ ($\rho = .318$, $p = .011$). A Fisher r -to- z comparison confirmed these two correlations did not differ, $z = 0.001$, $p = .999$. This indicates that the accuracy with which participants' predictions tracked their actual scores was essentially unchanged by AI assistance; what changed between conditions was not calibration accuracy but the direction and average magnitude of the miscalibration — participants were consistently optimistic in both conditions, and significantly more so when AI was used.

4.6 Qualitative Analysis of Open-Ended Questionnaire Responses

In addition to the quantitative results reported above, participants' open-ended responses to pre-task and post-writing questions were analyzed qualitatively. Duplicate or identical responses were counted only once per entry.

4.6.1 AI Tools Used by Participants

It was part of the design of this study that participants could use different AI writing tools, though they were requested to name the specific tool(s) they used. Table 1 below summarizes the frequency of use of the tools available in Saudi Arabia and that the participants could use individually or in combination ($N = 63$). This number exceeds 63 because many participants named more than one AI tool.

Table 1. Frequency of AI Writing Tools Used by Participants

Tool	Participants (n)
ChatGPT	54
Gemini	19
Grammarly	15
Google Translate	9
DeepSeek	4
Claude	2
Duolingo	1
Mindgrasp	1
QuillBot	1
Unclear / unclassifiable response	2

Note. $N = 63$; participants could name up to three tools.

Comparing different AI-tools, ChatGPT appears to be the most widely used technology—accounting for 86% of all responses, followed by Gemini (30%) and Grammarly (24%).

4.6.2 Typical AI-Assisted Writing Process (Pre-Task Reflection)

The results showed that the participants' responses fell into three common practice categories when they described their English writing processes. The most common approach—often referred to as the “post-writing” phase—involved writing independently first, using AI only to correct grammar, improve vocabulary or rewrite sentences (“I write what I want to say first, then use AI to make corrections later, not before”). A second group described an AI-assisted approach: using AI to research or organize ideas—especially when lacking a starting point—while personally handling the actual writing and structure (“I organize my thoughts, although not necessarily in a formal outline; I may use AI to generalize or identify ideas but, in the end, I draft the text myself”). A third smaller group described a hybrid (partly automated) approach: using AI tools to generate the bulk of the text, with the individual focusing mainly on final editing and revision (“using AI at every stage, from topic selection to final editing”). Participants described their use of AI in these scenarios—originally conceived in Arabic—using English terms.

4.6.3 Non-AI Condition: Process and Challenges

When describing the process of writing without AI tools, many participants revealed an unstructured approach: “writing whatever comes to mind without a specific plan” and “writing freely (without detailed planning) and editing the text later, if time permits.” These were common responses—highlighting an obvious contrast to the structured, goal-directed planning methods described earlier.

Several participants linked this directly to low confidence in their unassisted language ability (“I always feel my

wording is weak”; “I wrote only basic sentences to avoid grammatical mistakes because I don't trust my language level”). Notably, a subset of participants reported relying heavily on memorized content for exam-style writing and struggled specifically because this task used an unfamiliar topic with responses such as, “I can't write during exams unless I have already written about the topic before i.e., memorized it”; “I didn't plan before writing and felt I couldn't focus — this is a problem for me because I usually prefer planning, and in exams I rely on what I've already practiced.” The most frequently cited difficulties were spelling and grammatical errors, inability to generate or organize ideas for the given topic, self-doubt about their general language proficiency, and discomfort with not being able to use AI at all reflected in responses such as, “I've grown used to relying on it for writing about any topic.”

4.6.4 AI-Assisted Condition: Usage Patterns and Challenges

In describing how they used the AI tool during the AI-assisted task specifically, participants' responses again spanned from correction-only use (“I wrote it my way first, then used it for correction”) to full delegation (“I gave it the required topic and let it do the writing; all I did was ask it to shorten long sentences”), with several participants describing an iterative loop of requesting revisions until the output matched their intent (“I gave it the topic and let it write, then corrected what it wrote and kept editing or re-requesting edits until I reached the ideas I wanted to express”).

5. DISCUSSION

Results in the study indicate that though the AI intervention did not produce significantly different scores compared to non-AI-assisted writing, it contributed to refining the learners' writing, adjusting concepts, and producing content aligned with the writing objectives.

Taken together, findings in this study provide nuanced empirical support for the AI-assisted metacognitive disconnect described in the theoretical framework. The disconnect identified here is not a failure of AI to help, but a failure of self-assessment to track how much it helped: participants' predicted improvement exceeded their actual improvement to a significantly greater degree than the overestimation already present in their non-AI self-assessments. This pattern aligns closely with the affective-performance discrepancy reported by [7], in which AI-related gains in confidence and perceived competence do not scale proportionally with objectively measured gains, and with the regulatory gap described by [23], in which learners' metacognitive awareness fails to translate into accurate self-evaluation during AI-mediated writing.

The finding that within-condition calibration accuracy itself did not differ between AI-assisted and non-AI writing (identical correlations of $r = .34$ in both conditions) is an important qualification: participants were not systematically worse at judging their own work when AI was involved — they were, on average, similarly (imperfectly) calibrated in both conditions, but calibrated around a higher, more optimistic baseline when using AI. This aligns with the concept of [16] that AI-induced indirect reasoning processes and the framework of [19] on performance outcomes and judgment errors.

The weak correlation between non- versus AI-assisted performance ($r = .314$) further suggests that the level of outcome reproducibility reflects a stable individual characteristic, reminiscent of the verification paradox noted by [13]. Reliance on AI and its products can increase without a corresponding improvement in learners' ability to verify these results, meaning that reassessment in one context yields limited information about reproducibility in another context. This, along with the lack of association between general self-perception and actual performance, indicates that the observed difference in judgment is task- and context-specific rather than a stable feature of overconfident or underconfident individuals. Supporting the findings of [23], it is suggested that even students with self-assessment skills may have an inflated view of their abilities when working with AI in these scenarios.

In practice, these findings indicate that the primary risk associated with time-pressured AI-assisted writing is not the failure of the tool itself but rather the tendency of students to attribute more value to the AI output than is warranted. If not addressed, this may lead to a decrease in independent effort or an overreliance on technology. This is in line with the broader research on AI in education, which emphasizes the benefits of AI in delivering personalized feedback, detecting learning gaps, and the concerns about transparency, academic integrity, and over-reliance on AI systems [33]. This interpretation, along with its educational implications, is discussed in the results and recommendations sections below.

6. CONCLUSION

This study set out to examine the relationship between EFL learners' metacognitive awareness of their writing strengths and weaknesses and the quality of their AI-assisted versus non-AI-assisted writing under timed conditions. Using a quasi-experimental, repeated-measures design with 63 university-level EFL female learners, the study compared each participant's self-predicted and instructor-assigned rubric scores across both conditions.

Important conclusions follow directly from the results. First, AI assistance produced a genuine, objectively

measured improvement in writing quality under timed conditions: essays written with AI assistance scored higher on the instructor rubric ($M = 11.06/20$) than essays written without it ($M = 8.21/20$), a large effect ($d = 1.29$). This study therefore does not support a view of AI assistance as ineffective or harmful to writing quality in this population; on the contrary, it offers measurable support for grammatical, lexical, and organizational aspects of argumentative writing produced under time pressure.

Second, and central to the study's research question, learners' metacognitive awareness did not keep pace with this real improvement. Participants overestimated their own performance in both conditions, but the overestimation was significantly larger when AI was involved (MG-AI = 7.25 points) than without it (MG = 5.48 points), $t(62) = -4.50$, $p < .001$. The gap between predicted and actual scores widened by roughly two points on a 20-point scale specifically because of AI assistance, even though within-condition calibration accuracy — how well predictions tracked actual scores — did not differ between conditions ($r = .34$ in both cases). This is the clearest evidence in the dataset of an AI-assisted metacognitive disconnect: learners correctly perceived that AI helped them, but they systematically over-credited how much it helped.

Taken together, these findings refine rather than confirm the metacognitive disconnect construct as applied to AI-assisted EFL writing. The risk AI assistance poses under timed conditions, at least in this sample, is not that it fails to help, but that learners' confidence in its help exceeds the help itself — a distinction with direct implications for how AI tools should be introduced and scaffolded in the EFL writing classroom, discussed under Recommendations below.

7. RECOMMENDATIONS

The study proposes the following recommendations:

1. Integrating training on AI-supported task refinement:

Since the gap in understanding is created in AI-assisted scenarios, the refinement process should be embedded in the AI-supported task itself. For example, students can compare their initial drafts to the teacher's choices immediately after each assignment and provide detailed descriptions of which parts of the AI-assisted project they personally improved, as opposed to elements that were modified with minimal individual input.

2. Redesign assessment methods to reward meaningful engagement:

Programs that allow AI-assisted writing in structured or supervised environments should adjust their rubrics or employ process-based assessment. It encourages students to critically evaluate and refine AI-generated products rather than simply embrace them—especially given the important role AI plays in the final product.

3. Expand and diversify future samples:

Future research should apply this study design using larger, gender-balanced samples across different organizations, skill levels, topics, and narrative types to determine whether the observed differences extend beyond a single dialectical writing task.

4. Reporting scores for each level:

Researchers should record and report scores for each level of assessment—rather than aggregating scores—to allow for direct verification of the validity of the different levels.

5. Long-term intervention evaluation:

As discrepancies increase without a corresponding decrease in writing accuracy, future studies should investigate whether repeated practice can reduce students' overreliance on AI.

8. LIMITATIONS

The female only sample of the study limited to a single university reduces the broad reach of the findings to other, different learner samples. It is hoped that future endeavors will enroll more diverse participants to rule out this limitation.

REFERENCES

- [1] Al-Ahdal, A. A. M. H., & Hameed, P. F. M. 2026. EFL Learners' motivation to Write: A Study of the Efficacy of Quillbot and Grammarly. *Scientific Culture*, 12(1, Part 1), 2779.
- [2] Alkodimi, K.A., Al-Ahdal, A.A.M.H., 2025. Artificial Intelligence in Literature-Based Creative Writing: A Study with Story Live and AI Dun-geon. *Forum for Linguistic Studies*. 7(9): 754–769. DOI: <https://doi.org/10.30564/fls.v7i9.10615>
- [3] Alzabidi, A. S., & Al-Ahdal, A. A. M. H. 2025. E-Portfolios And Revision Logs: A Study Of Effects On Process Writing Approach In EFL. *Cultura: International Journal of Philosophy of Culture and Axiology*, 22(8s), 69–80.

- [4] Aslam, B., Nawaz, A.R., Ahmed, A.A., et al., 2025. Beyond Grammar Correction: Generative AI and the Cultivation of Higher-Order Thinking in EFL Writing. *Journal of Applied Linguistics and TESOL (JALT)*. 8(4), 388–404.
- [5] Azennoud, A., 2024. Enhancing Writing Accuracy and Complexity through AI-Assisted Tools among Moroccan EFL University Learners. *International Journal of Linguistics and Translation Studies*. 5(4), 211–226. DOI: <https://doi.org/10.36892/ijlts.v5i4.529>
- [6] Bjork, R.A., Dunlosky, J., Kornell, N., 2013. Self-Regulated Learning: Beliefs, Techniques, and Illusions. *Annual Review of Psychology*. 64(1), 417–444. DOI: <https://doi.org/10.1146/annurev-psych-113011-143823>
- [7] Chen, Y., Chen, J., 2025. Generative AI in EFL Writing Instruction for Lower-Intermediate Learners: Do Perceived Affective Gains Translate to Improved Performance? *Language Teaching Research*. 13621688251397724. DOI: <https://doi.org/10.1177/13621688251397724>
- [8] Costello, T.H., Newton, C., Lin, H., et al., 2023. A Metacognitive Blindspot in Intellectual Humility Measures. DOI: <https://doi.org/10.31234/osf.io/gux95>
- [9] De la Vall, F.O.N.T., Rebollo, R., Montoya, M.H., 2025. Retroalimentación Inteligente para Ensayos en Inglés como Lengua Extranjera. *International Journal of Foreign Languages*. (23), 1–26. DOI: <https://doi.org/10.17345/rile23.4174>
- [10] Elghani, Z., Asadzadeh-Maleki, N., 2026. The Effect of Artificial Intelligence Mediated Feedback on English Language Learners' Writing Ability. *Revista Metropolitana de Ciencias Aplicadas*. 9(1). DOI: <https://doi.org/10.62452/g6023m93>
- [11] Fernandes, D., Villa, S., Nicholls, S., et al., 2025. AI Makes You Smarter but None the Wiser: The Disconnect Between Performance and Metacognition. *Computers in Human Behavior*. 108779. DOI: <https://doi.org/10.48550/arXiv.2409.16708>
- [12] Guerrero-Chirinos, R.A., 2026. Generative AI Between Cognitive Offloading and Thinking: A Theoretical Framework of Metacognitive Tension in the Use of Generative AI. *Journal of Intelligent Decision Making and Information Science*. 3(4s), 546–557. DOI: <https://doi.org/10.59543/jidmis.v3.1042>
- [13] Huemmer, M., Durner, F., Shyiramunda, T., et al., 2026. AI, Metacognition, and the Verification Bottleneck: A Three-Wave Longitudinal Study of Human Problem-Solving. *arXiv preprint arXiv:2601.17055*. DOI: <https://doi.org/10.48550/arXiv.2601.17055>
- [14] Husnik, T., 2025. Metacognitive Strategies and the Use of ChatGPT in EFL Writing among Croatian High Schoolers [Doctoral dissertation]. Sveučilište u Zagrebu, Filozofski fakultet, Odsjek za anglistiku.
- [15] Kalamazh, R., Avhustiuk, M., 2018. Illusion of Knowing in Metacognitive Monitoring: Review of Possible Causes and Consequences. DOI: <https://doi.org/10.29038/2227-1376-2018-32-109-122>
- [16] Koch, 2026. Beyond the Steeper Curve: AI-Mediated Metacognitive Decoupling and the Limits of the Dunning-Kruger Metaphor. Available from: <https://arxiv.org/abs/2603.29681>
- [17] Koriat, A., Bjork, R.A., 2005. Illusions of Competence in Monitoring One's Knowledge During Study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 31(2), 187–194. DOI: <https://doi.org/10.1037/0278-7393.31.2.187>
- [18] Liu, Y., Wall, E., 2024. Unmasking Dunning-Kruger Effect in Visual Reasoning and Judgment. *IEEE Transactions on Visualization and Computer Graphics*. DOI: <https://doi.org/10.1109/tvcg.2024.3456326>
- [19] Maier, M., 2026. Self-Evaluation in AI-Assisted Cognition: An Explanatory Framework for Calibration and Miscalibration Effects. *Journal of Intelligence*. 14(7), 112. DOI: <https://doi.org/10.3390/jintelligence14070112>
- [20] Mekheimer, M., 2025. Generative AI-Assisted Feedback and EFL Writing: A Study on Proficiency, Revision Frequency and Writing Quality. *Discover Education*. DOI: <https://doi.org/10.1007/s44217-025-00602-7>
- [21] Muller, A.L., 2019. Neurological Correlates of the Dunning-Kruger Effect.
- [22] Mustafa, G., Shaikh, B., Raza, M.A., et al., 2026. Integrating Human-AI Collaborative Explainable Generative AI Feedback into EFL Writing Classrooms: Impacts on Academic Writing Performance, Learning Outcomes, Student Autonomy, and Learner Agency. *Journal of Applied Linguistics and TESOL (JALT)*. 9(1), 257–268.

- [23] Prather, J., Reeves, B.N., Leinonen, J., et al., 2024. The Widening Gap: The Benefits and Harms of Generative AI for Novice Programmers. In: Proceedings of the 2024 ACM Conference on International Computing Education Research—Volume 1, pp. 469–486. DOI: <https://doi.org/10.48550/arXiv.2405.17739>
- [24] RahmtAllah, E. A. E. 2025. The role of AI-powered chatbots in enhancing EFL learners' writing macro skills: Perceptions and challenges. *Forum for Linguistic Studies*, 7(12). <https://doi.org/10.30564/fls.v7i12.11937>
- [25] Rashid, S., Malik, S., Ghauri, F., 2026. The Active Learning-GenAI Synergy Framework: Ethical Integration of Generative AI in EFL/ESL Writing in Resource-Constrained Contexts. *F1000Research*. 14, 1210. DOI: <https://doi.org/10.12688/f1000research.171435.1>
- [26] Rian, I., 2025. AI-Calibrated Metacognition: How Genre-Based ChatGPT Feedback and Interaction Shape L2 Writers' Metacognitive Judgments and Self-Regulation. DOI: https://doi.org/10.31235/osf.io/pqzc9_v1
- [27] Stofiana, T., Sunendar, D., Mulyati, Y., et al., 2025. Writing with AI, Thinking with Toulmin: Metacognitive Gaps and the Rhetorical Limits of Argumentation. *Ampersand*. 100242. DOI: <https://doi.org/10.1016/j.amper.2025.100242>
- [28] Wang, P., Liu, T., Yang, Y., et al., 2025. Optimizing Self-Regulated Learning: A Mixed-Methods Study on GAI's Impact on Undergraduate Task Strategies and Metacognition. *British Journal of Educational Technology*. DOI: <https://doi.org/10.1111/bjet.70018>
- [29] Wang, X., Hussin, S., Aziz, A.A., 2025. Exploring the Effects of an Automated Writing Evaluation Tool on Metacognitive Engagement in Persuasive Writing. *International Journal of Learning, Teaching and Educational Research*. 24(10), 624–643. DOI: <https://doi.org/10.26803/ijlter.24.10.30>
- [30] Wilcox, J., 2022. How We Evaluate Our Thinking: The Accuracy of Our Metacognition. In: *Human Judgment*. SpringerBriefs in Psychology. Springer: Cham. DOI: https://doi.org/10.1007/978-3-031-19205-0_4
- [31] D. V. R. Moorthy, N. Anuradha, G. K. Devi, V. V. S. Rao, and C. Nagamani, "Education 5.0: Integrating artificial intelligence (AI) for personalized and adaptive learning," *Journal of Animal Environment*, vol. 18, no. 7S, pp. 631–637, 2026. <https://aejournals.org/index.php/AEJ/article/view/3300>
- [32] M. Kaliappan, S. Vimal, G. Dhiman, and K. Saranya, "Enhancing academic tutoring: A conceptual framework and preliminary evaluation of an AI platform leveraging large language models for content generation and knowledge augmentation," *Journal of Animal Environment*, vol. 18, no. 7S, pp. 638–646, 2026. <https://www.aejournals.org/index.php/AEJ/article/view/3301>
- [33] J. P. Pramod, P. G. Dharmik, N. Suresh, M. Vasantha Rao, and K. S. Chandu, "Education 4.0: Integrating artificial intelligence (AI) for personalized and adaptive learning," *Journal of Animal Environment*, vol. 18, no. 7S, pp. 75–81, 2026. <https://www.aejournals.org/index.php/AEJ/article/view/3128>
- [34] Pandey, P. (2025). Artificial intelligence for early detection of dental caries from radiographic images: A machine learning. *Journal of Dentistry and Clinical Sciences*, 1(1), 31–39. <https://agrpjournals.com/journals/journal/index.php/jdacs/article/view/12>
- [35] Malik, N., & Singh, K. (2025). Enhancing agro-industrial development in Haryana through crop diversification and value chain integration. *Journal of Artificial Intelligence in Agriculture and Applied Research*, 1(1), 13–19. <https://agrpjournals.com/journals/journal/index.php/jaiaar/article/view/4>