Journal of English Teaching, Literature, and Applied Linguistics
ISSN 2614-5871
Vol. 10, No. 2; 2026
Published by English Language Education Department of UMG
http://dx.doi.org/10.30587/jetlal.v4i1.128546

Exploring the Dunning-Kruger Effect in AI-Assisted EFL Writing among University Students

Najlah
English Department, Faculty of Teacher Training and Education
Universitas Muhammadiyah Palu
najlah13na@gmail.com

Ernitasari Mulyadi
English Department, Faculty of Teacher Training and Education
Universitas Muhammadiyah Palu
ernitasarimulyadi@mail.com*

Ismail Manangkari
English Department, Faculty of Teacher Training and Education
Universitas Muhammadiyah Palu
mailmanangkari@gmail.com

Arjuna Januari Monday
English Department, Faculty of Teacher Training and Education
Universitas Muhammadiyah Palu
junamndy2@gmail.com

Received:1st August 2026 Accepted: 1st September 2026 Published: 30th Septeber 2026

Abstract

The growing dependence of AI writing tools including ChatGPT, Grammarly, Gemini, QuillBot, and DeepSeek among university EFL learners increases an essential yet often overlooked question: are students actually getting better at writing, or are they merely getting better at using AI tools? This study explored that question by examining the role of the Dunning-Kruger Effect in the writing experiences of ten EFL university students at Universitas Muhammadiyah Palu who regularly use AI tools in their academic work. Data were collected through written asynchronous interviews and analyzed using Interpretive Phenomenological Analysis (IPA). Four themes emerged from the data: students showing confidence that outpaced their actual independent ability; students struggling to acknowledge their own contributions from AI-generated ones; a visible gap between how students assessed themselves and what they could actually do on their own; and perhaps most strikingly, students who recognized their over-reliance on AI but made no real effort to change it. Collectively, these findings suggest that AI tools affect more than the way students write. They gradually change students' views of their own writing skills, often in ways that do not match their actual writing ability. The study offers practical insights for EFL teachers, curriculum developers, and university administrators thinking through how AI should and should not be integrated into academic writing instruction.

Keywords: Dunning-Kruger Effect, metacognitive illusion, AI-assisted writing, EFL learners, Interpretive Phenomenological Analysis

1.Introduction

Authors retain copyright and grant the journal right of first publication with the work simultaneously licensed under a Creative Commons Attribution 4.0 International License **English Education Program Universitas Muhammadiyah Gresik, East Java, Indonesia**

There is something quietly unsettling happening in university EFL classrooms. Students are submitting better-looking writing assignments than ever before. Their grammar, vocabulary and ideas seem to develop almost in a night. But ask them to write something on their own without access to Artificial Intelligence tools and the gap between what they produce and what they submit can be surprising. This is not a story about cheating. It is a story about something more subtle and, in many ways, more concerning: the way AI writing tools are reshaping how students understand their own abilities.

The use of AI tools in academic writing has become almost universal among university students in Indonesia and beyond. Platforms like ChatGPT, Gemini, QuillBot, Grammarly, and DeepSeek are no longer uncommon they are part of the standard writing workflow for many learners. At one level, this makes sense. These tools are free, fast, and genuinely useful. They detect errors, suggest better phrasing, and help non-native speakers produce writing that sounds natural in English. For EFL learners navigating the demands of university-level academic writing in a second language, the appeal is obvious. Supporting other research revealed a hierarchy of AI tool preferences, with ChatGPT, Grammarly, and QuillBot being the most frequently used particularly during the editing and revision stages of writing. This supports the claim that these platforms have become a standard part of the writing workflow (Rofikah et al,. 2025). finding is further reinforced by (Arrozi et al., 2026), (Han et al., 2023) who found that students receiving ChatGPT-assisted instruction achieved significantly higher grammar and vocabulary scores compared to the control group. Similarly, (Han et al., 2023) reported that most students had positive experiences using ChatGPT for general, academic, and essay-writing purposes.

But there is a cognitive cost that tends to go unnoticed. When AI takes over the hard parts of writing students are relieved of exactly the kind of mental effort that produces genuine learning. Writing when done seriously is not just a way of recording ideas. It is a way of sharpening them. The struggle to find the right word. The effort to untangle a confusing sentence. The work of recognizing why a paragraph does not quite work. None of these are obstacles to writing. They are the process by which writers actually improve. When AI removes that struggle it also removes the learning embedded in it.

What makes this both interesting and concerning is what happens to students' self-perception in the process. Rather than recognizing that AI is doing cognitive work on their behalf, many students appear to interpret the quality of their AI-assisted output for the product of their own growing competence. This is precisely the dynamic described by the Dunning-Kruger Effect (Veiga et al., 2024): a situation where limited awareness of one's own limitations produces inflated confidence. In the AI-assisted writing context, that limited awareness is not accidental, instead it is manufactured by a tool that consistently delivers polished output regardless of the learner's actual skill level. Similarly, (Fernandes et al., 2026) people are unable to accurately assess their performance when using AI. Studies on overreliance in writing tasks define it as the acceptance of AI suggestions even when they are incorrect, often resulting in more erroneous decisions than would occur when working without AI assistance (Welzel & Vincent, 2026).

This study takes that problem seriously. Through written asynchronous interviews with ten university EFL students at Universitas Muhammadiyah Palu, it examines how the Dunning-Kruger Effect manifests in the lived experiences of learners who rely on AI tools in their academic writing. Rather than asking whether AI helps or harms writing competence a question that has been studied elsewhere this study asks something different: what does AI do to how students understand themselves as writers? The answers, drawn from participants' own words and reflections, reveal a pattern that has significant implications for how EFL writing is taught and assessed in the age of generative AI.

2.Method

This study employed a qualitative research design, specifically Interpretive Phenomenological Analysis (IPA)**,** to explore how university EFL students who regularly use AI writing tools make sense of their own writing ability and independence. IPA was selected because it is particularly suited to research questions concerned with how individuals interpret and construct meaning from significant personal experiences (Eatough & Smith, 2009).

Participants and Procedures

Ten university EFL students from the English Education program at Universitas Muhammadiyah Palu took part in this study. They were selected randomly from students who fit certain criteria relevant to the research questions. These criteria were current enrollment in a university English writing course regular use of at least one AI writing tool and a willingness to reflect honestly on their writing experiences. The sample included students from Semesters 2 4 and 6 with self-reported proficiency levels ranging from beginner (A2) to advanced. All ten regularly used AI tools though the specific tools and frequency of use varied. Table 1 presents their profiles.

Table 1. Participant Profiles

Code	Faculty/Major	Semester	Proficiency	AI Tools Used
S1	FKIP / English Education	6	Advanced	ChatGPT, BlackBox

Code	Faculty/Major	Semester	Proficiency	AI Tools Used
S2	FKIP / English Education	6	Intermediate	ChatGPT, Gemini
S3	FKIP / English Education	4	Intermediate	ChatGPT
S4	English Education	4	Beginner (A2)	ChatGPT, Claude AI
S5	FKIP / English Education	2	Intermediate	ChatGPT, Gemini
S6	FKIP / English Education	2	Intermediate	ChatGPT
S7	English Education	4	Beginner	ChatGPT, Grammarly
S8	FKIP / English Education	2	Beginner	ChatGPT, Gemini, DeepSeek
S9	English Department	4	Intermediate	ChatGPT, Grammarly
S10	English Education	6	Intermediate	ChatGPT

Data were collected through written asynchronous interviews. This method asks participants to respond to a set of structured questions in writing at their own pace over several days. This format was chosen on purpose. Metacognitive reflection does not happen quickly. It requires time and space for honest self-examination. Giving participants several days to respond rather than asking for immediate verbal answers created conditions where deeper more considered responses were possible.

The responses were then analyzed using Interpretive Phenomenological Analysis (IPA) a qualitative methodology that seeks to understand how individuals make sense of their own experiences (Smith et al. 2009). IPA fits this study particularly well because the research question is not about what AI does to writing quality in general. It is about what AI does to how specific individuals understand themselves as writers.

3. Findings and Discussion

Thinking About Thinking : The Theoritical Stakes

To understand what is happening when students over-rely on AI in writing, it is helpful to begin with examining metacognition itself — the capacity to think about one's own thinking. Flavell ((Stephanou & Mpiontini, 2017)), who introduced the term, identified two dimensions that matter greatly for writing. The first is metacognitive knowledge, or what a learner understands about themselves, the task at hand, and the strategies available to them. The second is metacognitive monitoring, the ongoing process of evaluating and adjusting one's thinking while a task is underway. Supported by other research revealed that metacognitive knowledge is defined as a fundamental component of metacognition and a prerequisite for self-regulated learning, comprising three types of knowledge: a. Person knowledge about oneself as a learner or cognitive processor; b. Task knowledge about the nature and demands of the task at hand, and c. Strategy knowledge knowledge about which strategies are effective for achieving specific cognitive goals (Zhang & Zhang, 2019). Other research revealed that Generative AI can effectively engage in human-like metacognitive processing without conscious awareness (Pavlović, 2024).

Good writers do both constantly. They notice when a sentence sounds off then figure out which parts need fixing. They recognize when their argument has drifted and pull it back. They know, in other words, what they know and what they do not. (Aykaç, 2026) stated that self-regulated learning model that views self-regulation as a skill developed through repeated practice within a supportive instructional environment (Jin et al., 2025).

AI interrupts this process in a specific and underappreciated way. When a student pastes a paragraph into ChatGPT and receives a revised version, the metacognitive work detecting the problem, diagnosing it, generating a solution, evaluating the result has been outsourced entirely. The student receives the product of that cognitive process without having to engage in the process itself. Over time, this means that the mental muscles required for self-monitoring and self-correction simply do not get exercised. The writing improves on the surface while the writer does not.

This is where the Dunning-Kruger Effect enters the picture. Kruger and Dunning (Veiga et al., 2024) demonstrated that people with limited competence in a domain tend to overestimate their own ability not because they are arrogant, but because the same deficits that make them perform poorly also prevent them from recognizing their poor performance. The same thing happens with AI writing. Leaning on AI to do the hard work stops students' independent writing skills from developing. It also stops them from noticing that gap because their submitted work always looks competent. So a learner ends up feeling like they are improving since their essays keep getting better. But they don't register why. The essays are getting better because the AI is getting better at helping them not because the student is actually growing. (Jin et al., 2025) found that as AI works harder to assist, the actual differences in ability between students become obscured in their output. This is precisely the mechanism you described: the work still looks good because AI bridges the gap, not because the students are growing. (Potts & Sudhof, 2026), novice users experience far more invisible failures conversations that appear to have been successfully concluded but actually miss the mark. For the

minimal-proficiency user category, 85.6% of their failures were invisible, compared to only 22.1% in the high-fluency group.

Zimmerman's (Fan & Gao, 2026) self-regulated learning (SRL) model adds another layer here. Effective learning moves through three phases. First comes forethought where a writer plans and sets goals. Then comes performance where the writer actually works and checks their own progress. Finally comes self-reflection where the writer looks back at the outcome and adjusts their strategy for next time. Each phase only works if the writer has accurate self-knowledge to build on. Similarly, (Elly & Rahmat, 2026) revealed that the use of AI for metacognitive scaffolding (AIM-SU) has the potential to enhance writing monitoring and regulation (MMRW), which aligns directly with Zimmerman's performance phase.

The problem is that AI often takes over the performance phase. It monitors for errors and makes corrections instead of the student. That leaves the self-reflection phase with nothing real to work with because students cannot honestly evaluate a process they did not actually go through themselves. And once self-reflection breaks down the next round of forethought breaks down too. The planning the goal-setting the strategic choices for the next writing task all become guesswork. The whole learning cycle quietly falls apart.

Finally (Gilbert & Risko , 2016) work on cognitive offloading explains why this matters beyond just one writing task. Offloading mental work onto outside tools makes sense in principle. Humans have always used tools to stretch their thinking further. The problem starts when offloading becomes so routine that the line between what the person can actually do and what the tool does for them starts to blur. Once that happens learners may genuinely start believing the tool's abilities are their own. This kind of misattribution feels natural to the mind but is dangerous for learning. Table 2 summarizes how each theoretical framework maps onto the AI-writing problem examined in this study.

Table 2. Theoretical Frameworks and Their Roles in This Study

Framework	Key Scholar(s)	Relevance to AI-Assisted Writing
Dunning-Kruger Effect	Kruger & Dunning	Explains how limited metacognitive awareness produces inflated self-confidence in AI-assisted writers
Metacognition Model	Flavell	Identifies the breakdown in monitoring and self-evaluation caused by AI substitution
Self-Regulated Learning	Zimmerman	Shows how AI disrupts the forethought-performance-reflection learning cycle
Cognitive Offloading	Risko & Gilbert	Explains how habitual AI use blurs the boundary between personal skill and tool capability

What the Data Revealed: Four Themes

Analysis of the ten interview responses produced four superordinate themes that together tell a coherent and somewhat sobering story about AI dependency and self-perception among university EFL learners.

Theme 1: Confidence That Runs Ahead of Competence

The most consistent pattern across the dataset was this. Participants reported growing confidence in their writing ability since they began using AI but could not point to evidence of genuine independent improvement. Their confidence was real. It was just attached to the wrong thing. It was confidence in their AI-assisted output not in their own developing skill. (Fisher & Oppenheimer, 2021); (Bouzar et al., 2024) revealed that when humans interact with AI systems, with individuals overestimating their own abilities even when the AI actually hinders their performance; this poor calibration to ambiguity regarding who deserves credit for success (or blame for failure) when external input is involved.

S10 put this with unusual directness. She acknowledged that her writing had become neater and more polished since using ChatGPT but immediately followed this with: "over time I have become dependent on AI and I feel that is not natural for me." When asked whether her submitted work represented her true ability her answer was simply: "No, because I use AI." There is something striking about this response. S10 is not confused about what is happening. She can see the gap between her AI-assisted work and her independent ability quite clearly. And yet as the rest of her interview made evident this clarity had not led her to change her behavior. She kept using AI kept submitting work that did not reflect her true ability and kept experiencing the confidence boost that came with producing polished essays even while knowing that the polish was not really hers.

S3 described a similar experience: "I feel good and happy because the quality looks cleaner and more professional. However, I feel it only partially represents my true ability because the AI helped me clean up the errors." The emotional register here feeling good and happy about AI-assisted work is worth noting. The confidence boost is real

on an emotional level even when the student knows it is not fully earned. This may be the most insidious part of AI-induced overconfidence. It feels genuine even when the student has reason to suspect it is not.

S4 a beginner-level student in Semester 4 framed the uncertainty in terms of attribution: "I feel like my writing ability has slightly improved since I started using AI tools, because they help me see better ways of expressing my ideas. However, I'm not fully sure how much of that improvement is really my own skill versus just getting used to AI's suggestions." This uncertainty this inability to cleanly separate personal growth from AI assistance is exactly what the Dunning-Kruger model predicts. Limited metacognitive awareness makes it hard to accurately locate the source of one's own performance.

Theme 2: Whose Writing Is This, Anyway?

The second theme that emerged from the data concerns authorial ownership. In simple words, it is about how hard it became for participants to know which parts of their writing were really theirs and which parts came from AI. This matters a lot. If a student cannot tell apart their own work from the AI's work they cannot know their own strengths and weaknesses. They cannot set clear goals for improving. And they cannot honestly say that the work they submit is fully their own a fact that most participants seemed to understand even if they did not say it directly. (Draxler et al., 2024), the finding that AI Ghostwriter Effect: users do not view themselves as the owners and authors of AI-generated text, yet they are simultaneously reluctant to publicly acknowledge AI as the author. In other words, they neither claim full ownership nor disclose AI's involvement creating a gray area of authorship.

S4 was the most honest about this confusion: "Honestly, it is often difficult for me. I frequently get confused about which parts are originally mine and which parts came from the AI, especially after I have revised something several times using its suggestions." The more a student revises with AI help the worse this confusion gets. Each round of AI suggestions makes the line between her ideas and the AI's ideas harder to see until she can no longer remember which parts she actually wrote herself.

S8 gave a response that was confusing in one way but also showed strong self-awareness in another way. Her English is still developing. Her interview answers have grammar mistakes that her AI-assisted essays probably do not have. But she has a clear sense of what AI changes about her voice. She gave an example to show the difference. Her own writing might sound like "there is never late time to stop learning unless its Doom day" while AI would change this to "it is never too late to learn, unless it is Judgment Day." She said: "Its still hard for me to form word that sound natural in English." What is interesting here is that S8 can actually notice the difference between her own voice and the AI's voice. She knows what AI is doing to her writing. But she still accepts these changes because the AI version sounds better. She is stuck between two good things she wants: sounding like herself and writing well. In her case AI wins almost every time.

S5 showed a clear understanding of this quality gap: "AI-generated writing is usually more organized, more professional, and has fewer grammar mistakes. My own writing is different because I am still learning, so I know that my sentences are not always as clear or as well-structured as the ones written by AI." This shows S5 is honest with herself. She can see the gap between her writing and AI writing. But it also shows a passive attitude toward using AI. The AI keeps producing better writing and the student keeps simply accepting it. There is little sign that she questions or challenges the AI's suggestions in a way that would turn AI use into real learning.

Theme 3: The Gap Between Self-Image and Reality

The third theme is maybe the clearest example of the Dunning-Kruger Effect in the whole dataset. It shows a gap that kept appearing between how participants rated their own writing ability and the limitations they admitted to at the same time. This is not just about students being wrong about their skills. It is about students holding two opposite beliefs at once. On one hand they rate themselves fairly positively. On the other hand they know they still have real weaknesses in independent writing. And most of them did not seem to notice this contradiction clearly. (Pennycook et al., 2017); (Pazicni & Bauer, 2013) examining the Dunning-Kruger effect in high-level reasoning, finding significant overestimation by participants with the lowest cognitive test scores; highlighting the role of metacognitive monitoring in recognizing one's own biases and errors.

S9 shows this pattern very clearly. She rated her writing ability at "around 7 out of 10" which sounds fairly confident. But in the same interview she admitted: "I still worry about using the right academic words and developing my ideas clearly. Sometimes I also feel my writing is still too simple." When asked directly whether her confidence matched her real ability she said: "Not always. AI makes me feel more confident because it improves my writing, but I know I still need more practice without it." The phrase "not always" says a lot here. It shows she partly realizes something is off in how she sees herself. But it stops far short of the kind of honest self-check that would actually fix this gap.

S6 shows an even stronger version of this problem. She said her confidence matched her real ability and gave this reason: "many of my friends and other people can understand what I write." This is an interesting standard to use. She is judging her writing by whether people can understand it not by whether it is actually correct or well written. This is an easy bar to pass even with real grammar and style problems especially since AI has already cleaned up her writing before anyone reads it. So her way of judging herself is based on the wrong measure. She is really judging AI-assisted writing but crediting herself for it.

S7's answer follows the same pattern the Dunning-Kruger model predicts for learners at a lower proficiency level. She described her writing as "good enough" and said it improved since she started using AI. But she also said she still worries about grammar and word choice and that her confidence is "not always" accurate. Having "good enough" and "still worried about grammar" in the same interview seems contradictory. But this is exactly the kind of confused self-assessment the Dunning-Kruger model expects from learners who do not yet have the tools to clearly judge their own performance. Table 3 maps the self-assessment patterns observed across all ten participants.

Table 3. Self-Assessment Patterns Across Participants

Code	Proficiency	Expressed Confidence	Acknowledged Limitation	Calibration
S1	Advanced	Confident writing independently	Relies on AI heavily for drafting	Overestimated
S2	Intermediate	Feels writing has improved	Needs more skill development	Partially accurate
S3	Intermediate	Happy with AI-assisted output	Work only partially reflects true ability	Partially accurate
S4	Beginner (A2)	Slightly improved; source unclear	Independent ability remains limited	Inaccurate
S5	Intermediate	Some improvement noted	Prefers learning independently	Partially accurate
S6	Intermediate	Confidence matches ability	Writing is simpler without AI	Overestimated
S7	Beginner	Writing is good enough	Still struggles with grammar and words	Overestimated
S8	Beginner	Confident but aware of insecurity	Fears AI may dull brainstorming	Partially accurate
S9	Intermediate	Rates self 7/10	Needs more practice without AI	Partially accurate
S10	Intermediate	Aware of dependency	Writing not polished without AI	Inaccurate

Theme 4: Knowing But Not Changing

The fourth theme was the most unexpected one and in many respects, the most significant for the theoretical framework of this study. Several participants clearly understood their own dependency on AI. They knew what was happening. They knew why it was a problem. They could even explain it in precise words. And yet this understanding had not changed how they actually behaved. They kept using AI in exactly the same way with the same frequency and the same passive habits even after describing the problem quite accurately. (Yeung, et al., 2026) revealed that, among users who misuse ChatGPT and are aware of the issue of AI hallucinations, some exhibit a paradoxical adoption pattern: they continue to rely on ChatGPT despite recognizing that it creates a new burden—the need to cross-check the results.

S4's closing reflection was especially striking in this regard:

> *I think the biggest thing I have realized through these questions is that I have been relying on AI more than I noticed, and it has made me want to be more mindful about using it to actually improve my writing, not just to make my writing look better. (S4)*

What makes this response interesting is its own meta quality. S4 is reflecting on what she learned about herself through the act of being interviewed. It was the interview itself not any instruction or feedback from a teacher that brought this insight forward. This suggests that metacognitive awareness of AI dependency may already exist quietly in many students. It is there but not activated. And structured chances for reflection like this interview could bring it to the surface in ways that actually matter.

S8 raised a concern about longer-term effects on thinking that shows real depth about the risks of using AI too much: "I still feel unsure for my own writing even though have using AI because sometime the AI is not able to really understand what i am about to say so its concerning. More, i think if use AI too much can make my brainstorming duller." This idea that regular AI use might slowly weaken her ability to think independently actually lines up well with

the cognitive offloading research (Gilbert, 2016). S8 reached a theoretically sound insight purely through her own experience without ever having read the academic frameworks that could explain it formally.

S10's answer about writing without AI was maybe the most emotionally honest in the whole dataset: "It's sad, because without AI, I feel my writing isn't polished." This sadness is real and worth taking seriously. S10 has built an identity as a writer that cannot be separated from her access to AI. Without it she feels not just less capable but somehow smaller as if the writing she can produce on her own does not even count as real writing. This is a deep kind of dependency and no amount of talking about academic integrity alone is likely to fix it. What is needed is something more. A teaching approach that helps students reconnect with the value and the dignity of their own unpolished developing and genuinely human writing.

Does Proficiency Level Make a Difference?

One of the study's secondary questions was whether proficiency level changes how the Dunning-Kruger pattern shows up among AI-using EFL writers. The data suggest a qualified yes. Proficiency level shapes the form the effect takes even if it does not decide whether the effect shows up at all.

Beginner-level participants (S4,S7,S8) tended to show stronger overconfidence compared to the limitations they admitted. This matches the original Dunning-Kruger finding. Lower competence tends to come with less accurate self-assessment because the metacognitive tools needed for accurate evaluation are themselves still underdeveloped. In the AI-assisted writing context this effect gets even bigger. Beginner-level students do not just lack the language skill to judge their writing accurately. They also lack enough familiarity with what good academic writing actually looks like to properly judge the gap between their AI-assisted work and their own independent work. stated that the relationship between miscalibration and self-confidence is negative for high performers but positive for low performers—specifically, the more miscalibrated low performers' self-estimates are, the more confident they are that those estimates are accurate. This supports the claim that awareness of what one knows and does not know depends on how much one already knows—precisely your argument regarding undeveloped metacognitive tools.

Intermediate-level participants generally gave more nuanced self-assessments. Their answers were more qualified more specific to context and more willing to admit particular weaknesses. But nuance is not the same thing as accuracy. Many intermediate participants admitted their limitations in general terms while still overestimating how much AI use had actually contributed to their real independent improvement.

The most interesting outlier in the dataset was S1 the only participant who identified herself as advanced. She expressed the strongest confidence in writing independently saying "I am confident that I can still write on my own without AI" while also describing a writing process fully built around AI for brainstorming drafting and refining. Her case is a reminder that higher proficiency does not protect someone from AI-induced metacognitive illusion. It may just create a more articulate and more convincing version of it.

On the other hand the most metacognitively accurate participant in the dataset was arguably S5 an intermediate-level student in Semester 2. Instead of using AI to generate or polish his writing he described using it mainly to understand the material before writing on his own. He acknowledged improvement "to some extent" a noticeably careful and limited claim and said he preferred learning through methods outside of AI. His relationship with AI tools is more strategic and less dependent than his peers and his self-assessment shows an accuracy that stands out clearly in the dataset.

What This Means for Teaching EFL Writing

The findings of this study point to several practical implications for EFL writing instruction that go beyond the usual advice to "use AI responsibly."

First, the theme of knowing but not changing shows that metacognitive awareness alone is not enough to actually shift behavior. Students need real structured chances to act on that awareness and this is something teaching needs to provide not something students can just figure out alone. Writing courses that include regular unassisted writing tasks clear self-assessment activities and structured reflection on the difference between AI-assisted and independent work are likely to work better than courses that simply allow or ban AI use without dealing with the metacognitive side of the problem. (Wischgoll, 2016); (L. Han, 2024) found that training in text structure knowledge combined with self-monitoring of the writing process demonstrates a significant effect in short-term interventions.

Second, the attribution errors seen in the data suggest students need help building a more precise vocabulary for talking about their own writing. When students cannot clearly say what they can and cannot do on their own they cannot set meaningful learning goals or ask for the right kind of help. Tasks that ask students to annotate their own writing marking which sentences they feel confident about which ones they are unsure of and which ones were heavily changed by AI could be a powerful way to build this kind of metacognitive precision.

Finally, the emotional side of AI dependency seen in S10's sadness at the idea of writing without AI and in S8's worry about her brainstorming becoming duller shows that teaching responses need to deal with more than just thinking habits. They need to deal with feelings too. Students who have come to feel that their unassisted writing is not good enough are unlikely to willingly practice it. Building a classroom where imperfect developing and unpolished writing is valued not just as a step toward better writing but as proof of real learning may matter just as much as any specific

teaching technique. (Zhang, 2026) revealed that writing anxiety is conceptualized as a multidimensional construct comprising cognitive concerns, somatic tension, and avoidance behaviors.

4. Conclusion

This study set out to explore a question that sounds simple but really is not. When university EFL students use AI to help them write what happens to how they see themselves as writers? The answer comes from the real experiences of ten participants at Universitas Muhammadiyah Palu and it is both clear and a bit concerning. AI writing tools are not only changing what students produce. They are changing how students see themselves. And for most participants this change moves toward overconfidence. Their sense of writing ability keeps growing but it is not built on real independent skill. It is built on the polish that AI keeps adding to their work.

Four themes together show the shape of this problem. Students feel confident in ways that go further than their actual independent ability. They struggle to clearly tell which parts of their writing are truly their own. They hold a positive view of themselves while also admitting real limitations without fully noticing that these two things do not match. And maybe the most telling part several participants clearly understood their own dependency on AI and still changed nothing about how they write. This last point deserves the most attention from teachers because it shows that the problem is not simply about students not knowing. Students can see what is happening to them. What they are missing is not awareness. What they are missing is real support that helps them act on what they already know.

The Dunning-Kruger Effect here is not just something happening inside one student's head. It is something built into how AI writing tools are currently used in EFL classrooms. These tools are extremely good at making student writing look competent and just as good at letting students avoid becoming competent themselves. Fixing this will take more than rules about when AI can or cannot be used. It will take a real rethink of what writing instruction is actually for. Not the production of polished essays but the growth of writers who can think clearly judge themselves honestly and improve on purpose whether or not a machine is helping them.

5. References

Arrozi, F., Sunarti, A. H. (2026). Journal of English Language Teaching in Indonesia. *Eltin journal:* 373–382.

Aykaç, N. (2026). *Enhancing Writing Self-Regulation and Metacognitive Awareness Through Structured Free Writing : A Mixed-Methods Study*. *Journal of Inteligence*, 14(7)

Bouzar, A., Idrissi, K. E. L., & Ghourdou, T. (2024). *ChatGPT and Academic Writing Self-Efficacy: Unveiling Correlations and Technological Dependency among Postgraduate Students*. Arab World English Journal (AWEJ), 225–236

Draxler, F., Werner, A., & Munich, L. M. U. (2024). *The AI Ghostwriter Effect : When Users do not Perceive Ownership of AI-Generated Text but Self-Declare as Authors*. ACM Transactions on Computer Human Interaction, *31*(2).

Eatough, V., & Smith, J. A. (2009). *Interpretative Phenomenological Analysis*. In C. Willing & W. Stainton Rogers (Eds), *The SAGE handbook of qualitatif research in psychology.* 179–194.

Elly, O., Rahmat, NH. (2026). *The Shared Scaffold: A Metacognitive Framework of AI's Influence on Self-Regulated Writing*. International Journal of Research and Innovation in Social Science. *10*(21), 1044-1048. **DOI:** 10.47772/IJRISS.2026.102100088

Fan, K., & Gao, T. (2026). Acta Psychologica The role of AI-mediated metacognitive interventions in EFL writing development : A cross-sectional study of textual accuracy , lexical sophistication , and cohesion. *Acta Psychologica*.. https://doi.org/10.1016/j.actpsy.2026.107069

Fernandes, D., Villa, S., Nicholls, S., Haavisto, O., & Buschek, D. (2026). *AI makes you smarte but none the wiser: The disconnect between performance and metacognition*. Computers in Human Behaviour Journal . https://doi.org/10.1016/j.chb.2025.108779

Fisher, M., & Oppenheimer, D. M. (2021). *Who Knows What ? Knowledge Misattribution in the Division of Cognitive Labor*. Journal of Experimental Psychology : Applied *27*(2), 292–306.

Gilbert, E. F. R. & S. J. (2016). *Cognitive Offloading. University College London, London, UK*. 1–37.

Han, J., Kim, Y., Kim, M., Lim, H., Kim, J., & Lee, T. Y. (2023). RECIPE : How to Integrate ChatGPT into EFL Writing Education. In *Proceedings of the Tenth ACM Conference on Learning @Scale.* Association for Computing Machinery. https://doi.org/10.1145/3573051.3596200

Han, L. (2024). *Metacognitive Writing Strategy Instruction in the EFL Context : Focus on Writing Performance and Motivation*. SAGE Open, 1–16. https://doi.org/10.1177/21582440241257081

Jin, Y., Yang, K., Martinez-maldonado, R., Gašević, D., & Yan, L. (2026). *The Agency Gap in AI-Supported Writing : How Reactive and Proactive Agent Designs Shape Multimodal Reasoning*. 1–30. Computers and Education: Artificial Intelligence *11*, 100655. https://doi.org/10.1016/j.caeai.2026.100655

Pavlović, J. (2024). *Generative AI as a metacognitive agent: A comparative mixed-method study with human participants on ICF-mimicking exam performance*. https://doi.org/10.48550/arXiv.2405.05285

Pazicni, S., & Bauer, C. F. (2013). *Characterizing illusions of competence in introductory chemistry students*. Chemistry Education Research and Practice. *15*(1), 24–34. https://doi.org/10.1039/c3rp00106g

Pennycook, G., Ross, R. M., Koehler, D. J., & Fugelsang, J. A. (2017). *Dunning – Kruger effects in reasoning : Theoretical implications of the failure to recognize incompetence*. Psychonomic Bulletin & Review, *24*(6),1774–1784. https://doi.org/10.3758/s13423-017-1242-7

Potts, C., & Sudhof, M. (2026). *A paradox of AI fluency*. *arXiv*. https://doi.org/10.48550/arXiv.2604.25905

Rofikah, U., & Arafah, B. (2025). *Indonesian Students' Perception of AI -Assisted EFL Academic Writing*. In the Proceedings of the 5th International Conference on Linguistics and Cultural Studies 5. Atlantis Press https://doi.org/10.2991/978-2-38476-394-8

Stephanou, G., & Mpiontini, M. (2017). *Metacognitive Knowledge and Metacognitive Regulation in Self-Regulatory Learning Style , and in Its Effects on Performance Expectation and Subsequent Performance across Diverse School Subjects*. Psychology, 8(12), 1941–1975. https://doi.org/10.4236/psych.2017.812125

Veiga, M., Coutinho, C., Thomas, J., Fredricks-lowman, I., & Alkaabi, S. (2024). *Unskilled and unaware : second-order judgments increase with miscalibration for low performers*. Frontiers in Psychology. *15 Article* 1252520. https://doi.org/10.3389/fpsyg.

Welzel, V. H. A., & Vincent, N. (2026). Overreliance in Writing Tasks : Exploring Similarity-Based Measures of AI Influence on Writing and Proposing a Reflective Writing Interface Intervention. *The 2026 ACM Conference on Fairness, Accountability, and Transparency*. *5941-5958*. Association for Computing Machinery. https://doi.org/10.1145/3805689.3812389

Wischgoll, A. (2016). *Combined Training of One Cognitive and One Metacognitive Strategy Improves Academic Writing Skills*. Frontiers in Psychology. 1–13. https://doi.org/10.3389/fpsyg.2016.00187

Yeung, Lorraine, Ka, C., Pui, D., Chow, L., Hang, P., Sam, W., & Shun, S. (2026). *In ChatGPT they trust : a study of students ' perceptions and misuse of ChatGPT in higher education*. AI and Ethics, *6*(11).https://doi.org/10.1007/s43681-025-00855-w

Zhang, D., & Zhang, L. J. (2019). *Metacognition and Self-Regulated Learning (SRL) in Second / Foreign Language Teaching*. In X. Gao (Ed.), *Second handbook of English language teaching*. Springer, 1-15. https://doi.org/10.1007/978-3-319-58542-0_47-1

Zhang, Y. (2026). *AI-associated academic writing anxiety in AI-assisted contexts : evidence from Chinese EFL postgraduate students*. Frontiers in Psychology, 1–14. https://doi.org/10.3389/fpsyg.2026.1841607