

From Silent Thinkers to Active WRITERS: AI Chatbots in EFL Idea Expression

Angel Jarbowah^{1*}, Nanik Retnowati², Nuraeni³

English Language Study Program

FKIP, Universitas Ibnu Khaldun Bogor, Indonesia

¹ angelmohni07@gmail.com, ² nanik.retnowati@uika.bogor.ac.id, ³ nuraeni@uika-bogor.ac.id

* corresponding author

Abstract

The growing presence of artificial intelligence in education has opened new possibilities for supporting English as a Foreign Language (EFL) writing, yet the specific role of AI chatbots in helping students express ideas remains underexplored. This study investigates how AI chatbots such as ChatGPT, Gemini, and Claude support idea expression among EFL university students, addressing the phenomenon of “silent thinkers” who learners rich in ideas but unable to translate them into written form. Using a mixed-methods design, primarily qualitative and supported by quantitative data, 30 EFL university students completed a Google Form questionnaire comprising 19 Likert-scale items and 6 open-ended questions. Quantitative data were analyzed descriptively, and qualitative data were analyzed thematically following Braun and Clarke’s framework. Findings, interpreted through Vygotsky’s Sociocultural Theory and Krashen’s Affective Filter Hypothesis, reveal that AI chatbots function as scaffolding tools that support brainstorming, idea generation, vocabulary access, and sentence construction, while also reducing writing anxiety and increasing confidence and engagement. Students nonetheless expressed concern about overdependence, plagiarism, and inaccurate AI output, and emphasized the necessity of teacher guidance. The study concludes that AI chatbots, used responsibly, can transform silent thinkers into active writers.

Keywords: *AI Chatbots, EFL Writing, Idea Expression, Scaffolding, Writing Confidence.*

INTRODUCTION

The digital transformation of education has fundamentally reshaped how teaching and learning occur between institutions worldwide. Over the past decade, digital tools that once served as supplementary aids by learning management systems, online quizzes, and recorded lectures have become central to instructional practice itself (Zawacki-Richter et al., 2019). Within this transformation, the emergence of

generative artificial intelligence (AI) and significantly Large Language Model (LLM)-based conversational tools which represents a development of a qualitatively different order. Where earlier educational technologies digitized existing practices, tools such as ChatGPT, Google Gemini, and Anthropic's Claude introduce a genuinely new kind of interlocutor into the learning process: one capable of producing fluent, contextually responsive language, sustaining extended dialogue, and adapting its output to the learner's immediate needs (Kasneci et al., 2023). Within language education, the uptake of these tools has been especially swift and especially significant.

For learners of English as a Foreign Language (EFL), writing has always represented one of the most cognitively and affectively demanding productive skills to develop. Unlike writing in one's native language, EFL writing demands the simultaneous management of vocabulary retrieval, grammatical encoding, organizational planning, and rhetorical awareness all under a linguistic system that has not yet been automatized through years of native exposure (Amoush & Alhosban, 2025). The result for many EFL learners is a chronic struggle not merely with the surface features of written English but with the deeper challenge of idea expression: the ability to generate, develop, and commit one's thoughts to the page in a language that does not yet feel fully one's own. It is this challenge in which is prior to grammar, prior to organization, prior to style that sits at the center of this study.

A growing and largely encouraging body of research has documented AI chatbots' capacity to address these challenges. Song and Song (2023) found that EFL university students who used ChatGPT in their writing process reported measurable improvements in writing skills alongside heightened motivation, suggesting that AI support operates simultaneously on cognitive and affective levels. Mahapatra (2024) documented specific writing performance gains in a structured ChatGPT-integrated ESL writing intervention. Duong and Chen (2025) found that more frequent and purposeful student engagement with an AI writing chatbot correlated with stronger writing outcomes and more positive attitudes. Slamet (2024) who worked significantly within the Indonesian EFL context highlighted that both teachers and students viewed ChatGPT favorably as a digital learning assistant, particularly valued for its round-the-clock availability and non-evaluative responsiveness. Escalante et al., (2023) found that AI-generated feedback was judged effective by a majority of EFL learners and preferred over peer feedback. Gao et al., (2025) reported a positive relationship between AI chatbot-integrated writing tasks and both writing self-efficacy and critical thinking. Yan

and Zhang (2024) traced how L2 writers engaged with ChatGPT's corrective feedback over time by finding patterns of strategic yet selective use rather than uncritical dependence. Wang (2025) who compared native and non-native English speakers, found that non-native speakers relied on AI tools heavily for idea generation in which is a finding that points directly toward the construct this study centers, not writing quality as a product yet the idea expression as a process.

Beneath and prior to all the linguistic and organizational dimensions of EFL writing lies a stage that receives insufficient dedicated attention, the movement from silence to articulation from unformed thought to committed written language. Many EFL learners are not failed writers in any technical sense, yet they are as this study terms them, "*silent thinkers*" – students who carry genuine ideas, reasoning, and perspectives but who remain unable to bridge the gap between internal thought and English text not for lack of grammatical knowledge but because the cognitive and affective labor of idea expression in a foreign language exceeds what they can sustain unsupported.

Despite the broadly positive picture sketched above a substantial body of scholarship has raised concerns that any responsible account of AI in EFL writing must take seriously. Cotton et al., (2024) underlined an influential cautionary analysis, by warning that the ease with which students can generate polished text using ChatGPT creates meaningful risks for academic integrity and risks allowing students to bypass the cognitive effort that writing instruction is designed to cultivate by producing the appearance of competence without its substance. Amoush and Alhosban (2025) in a systematic review focused specifically on ChatGPT in EFL writing, provided a broader set of risks: overreliance that gradually erodes students' independent writing capacity, the generation of factually inaccurate or contextually inappropriate content, and the displacement of students' own authorial voice by AI-generated phrasing. Jung et al., (2026) through applying Activity Theory to EFL writing classrooms where ChatGPT had been integrated, found that the tool's affordances came into structural tension with existing norms around assessment, academic integrity, and pedagogical goals which are tensions that cannot be resolved simply by permitting or prohibiting AI use but that require deliberate, sustained instructional redesign.

These concerns are real and important. Yet they do not constitute a reason to abandon the empirical investigation of AI chatbots' potential benefits, they constitute a reason to conduct that investigation more carefully, and with a clearer sense of what

specifically is being claimed. It is at precisely this point that the present study identifies its contribution. A close reading of the existing literature reveals a specific and consequential gap: the great majority of studies examining AI chatbots in EFL writing have framed their central outcomes in terms of writing quality, performance scores, self-efficacy, or even general perceptions of usefulness. Few have isolated idea expression itself which in the brainstorming, the thought organization, the initial movement from silence to articulation as a primary and theoretically grounded object of inquiry. This is a significant omission, because idea expression is the most foundational and most universally experienced obstacle in EFL writing. It precedes and underlies every subsequent dimension that writing research typically measures.

This gap becomes more pointed when examined through the theoretical lenses that the present study applies. Vygotsky's (1978) Sociocultural Theory, with its account of learning as dialogically scaffolded within the learner's Zone of Proximal Development (ZPD), would predict that AI chatbots by functioning as a responsive, always-available more knowledgeable other should be most powerfully effective precisely at the ideational stage, where scaffolding can extend a learner's capacity furthest beyond their independent baseline. Krashen's (1982) Affective Filter Hypothesis would similarly predict that the non-evaluative, patient quality of AI chatbots should lower the anxiety that prevents many EFL students from attempting to express ideas in writing at all. Both predictions remain largely untested with respect to idea expression specifically.

This study is designed to fill that gap. Its purpose is to investigate how AI chatbots support idea expression among EFL university students in writing activities, and to examine how students themselves perceive and experience that support by including both its benefits and its limitations. The study is guided by two research questions.

1. To what extent do AI chatbots facilitate idea expression by including brainstorming, idea generation, and thought organization among EFL university students during writing activities?
2. How do EFL university students perceive the role of AI chatbots in their writing classrooms, particularly regarding confidence, anxiety reduction, engagement, and responsible AI use?

The study contributes to the field in three ways. First, by centering idea expression as the primary construct rather than a secondary dimension of writing quality, it addresses a specific and previously underexamined gap in the AI-EFL writing

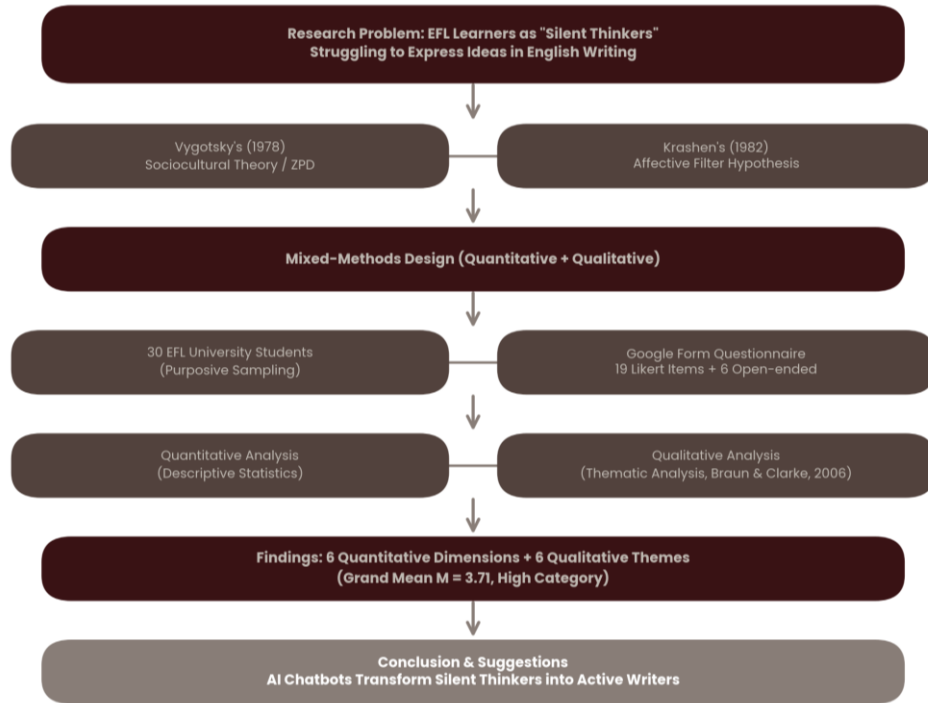
literature. Second, by applying Vygotsky's Sociocultural Theory and Krashen's Affective Filter Hypothesis in tandem, which is going to offer a theoretically integrated account of both the cognitive scaffolding and the affective filtering mechanisms through which AI chatbots may operate at the ideational stage of writing. Third, through a mixed-methods design involving 30 EFL university students mainly qualitative and supported by descriptive quantitative data in which it provides a richly student-voiced account of AI chatbot use that complements the more aggregative, product-focused findings of much existing research.

METHOD

Research Design

This study employed a mixed-methods design in which was primarily qualitative in orientation and supported by descriptive quantitative data. A mixed-methods approach was selected due to the study's central interest by how EFL students experience, interpret, and make sense of AI chatbot support during idea expression is fundamentally a question about lived experience and perception, best illuminated through students' own words. At the same time, quantitative data drawn from a structured Likert-scale instrument provided a systematic and comparable overview of the breadth and consistency of those perceptions across the full sample by allowing the qualitative themes to be read against a clear descriptive backdrop. The two data strands were collected concurrently through a single instrument and analyzed in a mutually informing manner: quantitative patterns identified where to look closely, and qualitative accounts explained why those patterns existed. Figure 1 presents an overview of the research design flow.

Figure 1. Overview of the Research Design Flow



Participants

The participants were thirty Indonesian EFL university students enrolled in English language education programs participated in this study. All participants had direct experience by integrating AI chatbots in their English writing activities and were selected through purposive sampling to maintain assurance that each respondent could speak from genuine lived experience with the phenomenon under investigation. Table 1 presents a complete summary of the participant profile, including gender, academic semester, primary AI chatbot tool used, and frequency of AI chatbot use in writing activities.

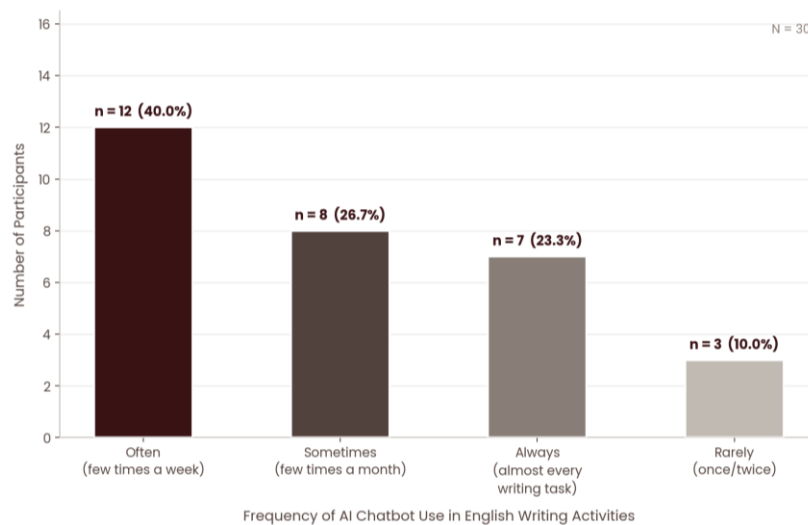
Table 1. Distribution of Participants Profile

Category	Group	n (%)
Gender	Female	20 (66.7)
	Male	10 (33.3)
Semester	1–2	6 (20.0)
	3–4	11 (36.7)
	5–6	11 (36.7)
	7–8	2 (6.7)

AI Tool Used	ChatGPT	20 (66.7)
	Google Gemini	8 (26.7)
	Claude	2 (6.7)

As shown in Table 1, ChatGPT was the dominant tool among participants, used by 66.7% of the sample, followed by Google Gemini (26.7%) and Claude (6.7%). This distribution is broadly consistent with global patterns of AI chatbot adoption in educational contexts. While the findings are most heavily anchored in ChatGPT-mediated writing experience, the inclusion of Gemini and Claude users ensures that the study also captures meaningful variation in platform-specific affordances. In the bellow figure which represents the majority of participants reported using AI chatbots often or very frequently in their writing activities, with 63.3% using them several times a week or more, indicating a relatively AI-engaged sample well positioned to reflect on the experience under investigation.

Figure 2. Distribution of Participants Reported AI Usage



Data Collection Instrument

Data were collected by utilizing a single Google Form questionnaire comprising two integrated components. The first component consisted of 19 Likert-scale items each rated on a five-point response scale (1 = Strongly Disagree until 5 = Strongly Agree), organized across six conceptually distinct dimensions. The second component consisted of six open-ended questions designed to elicit rich, free-response qualitative

data. Table 2 presents the full structure, dimensions, item coverage, and theoretical alignment of the instrument.

Table 2. Structure and Theoretical Alignment of the Questionnaire

Section	Dimension	Items	Focus	Theory
A: Profile	Respondent Background	3 items	Gender, semester, AI tool used	-
B: Likert Scale	1. <i>Idea Generation & Brainstorming</i>	Q1 – Q4	Generating ideas, overcoming writer's block, starting to write	Vygotsky (1978) – ZPD & Scaffolding
	2. <i>Vocabulary & Sentence Support</i>	Q5 – Q7	Word choice, sentence construction, linguistic precision	Vygotsky (1978) – ZPD & Scaffolding
	3. <i>Writing Confidence</i>	Q8 – Q10	Self-assurance, willingness to attempt complex tasks	Krashen (1982) – Affective Filter Hypothesis
	4. <i>Writing Anxiety</i>	Q11 – Q13	Nervousness, fear of mistakes, apprehension when writing	Krashen (1982) – Affective Filter Hypothesis
	5. <i>Engagement & Participation</i>	Q14 – Q16	Motivation, active participation, writing frequency	Krashen (1982) – Affective Filter Hypothesis
	6. <i>Overall Perception of AI</i>	Q17 – Q19	Global attitude toward AI-assisted EFL writing	Both frameworks
C: Open-ended	Qualitative Experience	6 items	Benefits, difficulties, ethics, overdependence, teacher guidance	Both frameworks

As shown in Table 2, the six Likert-scale dimensions were designed to move systematically from the most process-specific aspects of writing support regarding idea generation and vocabulary access through the affective dimensions of confidence, anxiety, and engagement, to a final integrative measure of overall perception. Theoretically, Dimensions 1 and 2 map most directly onto Vygotsky's (1978) scaffolding model, as they capture the cognitive and linguistic support AI chatbots provide within learners' Zones of Proximal Development. Dimensions 3, 4, and 5 map most directly onto Krashen's (1982) Affective Filter Hypothesis, capturing the emotional variables in which are confidence, anxiety, and motivation that the hypothesis identifies as central mediators of language production. Dimension 6 draws on both frameworks and provides a summary-level evaluative perspective. The six open-ended questions in

Section C served as the primary data source for the thematic analysis that forms the backbone of the Results and Discussion section.

Data Analysis

Quantitative data from the 19 Likert-scale items were analyzed using descriptive statistics: means, frequency counts, and percentage distributions which have been calculated separately for each item and aggregated by dimension. Mean scores were interpreted using the following five-interval scale: 1.00-1.80 = Very Low; 1.81-2.60 = Low; 2.61-3.40 = Moderate; 3.41-4.20 = High; 4.21-5.00 = Very High. Dimension-level means were computed by averaging the item means within each of the six dimensions, and an overall grand mean was computed across all 19 items.

Qualitative data from the six open-ended questions were analyzed using thematic analysis following Braun and Clarke's (2006) six-phase procedure:

1. Familiarization with the data through repeated careful reading of all 30 sets of responses.
2. Generation of initial codes by systematically labeling meaningful units of data.
3. Construction of candidate themes by collating related codes.
4. Review and refinement of themes against the full data set.
5. Clear definition and naming of each final theme.
6. Production of the written report with sufficient illustrative evidence.

Because the study's design was primarily qualitative, this thematic analysis formed the analytic backbone of the Results and Discussion section, with the quantitative dimension means providing a structured descriptive frame before each theme was elaborated in depth through participant voices. Data triangulation across both strands was employed throughout to enhance the credibility and completeness of the findings.

RESULTS AND DISCUSSION

This section presents and discusses the findings from the quantitative Likert-scale data and the qualitative open-ended responses. The 19 Likert-scale items across six dimensions are reported through descriptive statistics including means and frequency distributions. Mean scores are interpreted using the five-interval scale: 1.00-1.80 = Very Low; 1.81-2.60 = Low; 2.61-3.40 = Moderate; 3.41-4.20 = High; 4.21-5.00 = Very High. The open-ended responses were analyzed through thematic analysis (Braun

& Clarke, 2006), yielding six themes discussed in relation to Vygotsky's (1978) Sociocultural Theory and Krashen's (1982) Affective Filter Hypothesis.

Table 3 presents the item-level means and frequency distributions for all 19 items. The grand mean was $M = 3.71$ (High), reflecting consistently positive perceptions of AI chatbot support across all six dimensions. The highest-scoring dimension was Vocabulary and Sentence Support ($M = 3.89$), followed by Idea Generation and Brainstorming ($M = 3.88$). The single highest-scoring item was Q1 "AI chatbots help me generate ideas when I don't know what to write about" ($M = 3.97$), with 80.0% of participants responding Agree or Strongly Agree. The lowest-scoring dimension was Engagement and Participation ($M = 3.54$), though all six dimensions fell within the High category.

Table 3. Item-Level Mean Scores and Frequency Distributions (N = 30)

Item Statement	SD	D	N	A	SA	M	Cat.
Dimension A: Idea Generation & Brainstorming							
Q1. AI chatbots help me generate ideas when I don't know what to write about.	0	2	4	17	7	3.97	High
Q2. Using AI chatbots makes it easier for me to start writing.	0	0	8	16	6	3.93	High
Q3. AI chatbots help me organize my thoughts before writing.	1	0	7	17	5	3.83	High
Q4. AI chatbots give me useful suggestions when I experience writer's block.	0	1	11	11	7	3.80	High
Dimension Mean						3.88	High
Dimension B: Vocabulary & Sentence Support							
Q5. AI chatbots help me find the right words to express my ideas in English.	0	1	6	18	5	3.90	High
Q6. AI chatbots help me construct better sentences in my writing.	0	2	5	18	5	3.87	High
Q7. Using AI chatbots makes my written ideas clearer and more precise.	0	1	8	14	7	3.90	High
Dimension Mean						3.89	High
Dimension C: Writing Confidence							
Q8. I feel more confident writing in English when I use AI chatbots.	1	4	10	10	5	3.47	High
Q9. AI chatbots encourage me to express my ideas more freely in writing.	0	0	9	17	4	3.83	High
Q10. Using AI chatbots reduces my fear of making mistakes in writing.	1	3	9	13	4	3.53	High
Dimension Mean						3.61	High
Dimension D: Writing Anxiety							
Q11. I feel less anxious about writing in English when I have access to AI chatbots.	0	2	15	10	3	3.47	High
Q12. AI chatbots make the English writing process feel less stressful for me.	0	3	9	15	3	3.60	High
Q13. I feel more comfortable expressing my ideas in writing with AI chatbot support.	2	2	6	15	5	3.63	High

Dimension Mean							3.57	High
Dimension E: Engagement & Participation								
Q14. Using AI chatbots makes me more motivated to complete writing tasks.	1	4	10	11	4		3.43	High
Q15. I participate more actively in writing activities when I can use AI chatbots.	2	2	7	14	5		3.60	High
Q16. AI chatbots make English writing activities more enjoyable for me.	1	3	6	17	3		3.60	High
Dimension Mean							3.54	High
Dimension F: Overall Perception of AI Chatbots								
Q17. Overall, I think AI chatbots are helpful tools for EFL writing.	1	0	8	15	6		3.83	High
Q18. I would recommend AI chatbots to other EFL students for writing support.	1	1	11	11	6		3.67	High
Q19. I believe AI chatbots improve the quality of my written work.	2	0	6	18	4		3.73	High
Dimension Mean							3.74	High
Grand Mean (All 19 Items)							3.71	High

SD = Strongly Disagree; D = Disagree; N = Neutral; A = Agree; SA = Strongly Agree. M = Mean. Cat. = Category. Scale: 3.41–4.20 = High.

The thematic analysis of the open-ended responses yielded six themes that map directly onto the six questionnaire dimensions and together provide a rich qualitative account of the experiences and processes the quantitative data describe in aggregate. Each theme is presented and discussed below.

Theme 1: AI Chatbots as Brainstorming and Idea Generation Partners

The first and most prevalent theme (93.3% of responses) emerged from Dimension A (M = 3.88) and concerned AI chatbots' dual role as brainstorming partners and idea generation tools. Participants described turning to ChatGPT, Gemini, or Claude not after deciding what to write, but precisely when they did not yet know what to say by using the chatbot as a conversational sounding board that helped them move from a blank page to an initial cluster of ideas. Beyond idea generation, students also described using the chatbot to organize ideas already in mind but not yet structured, distinguishing clearly between not knowing what to write and knowing but not knowing how to arrange it:

"When I have no idea what to write, I ask the AI to give me several possible angles, and then I pick the one that resonates with me the most. It is like having a thinking partner who is always ready" – (R10)

"Sometimes I know what I want to say in my head but I don't know how to arrange it properly. I describe my ideas to the AI and ask it to help me make an outline. Then I write from the outline myself" – (R17)

These accounts map directly onto Vygotsky's (1978) concept of the Zone of Proximal Development: the chatbot functions as a more knowledgeable other that scaffolds the student's thinking into territory unreachable without assistance, without completing the task on the student's behalf. These findings align with Duong and Chen (2025), who found EFL students' AI engagement heaviest during pre-writing, and with Wang (2025) and Gao et al. (2025), who found AI tools supported both initial idea production and the higher-order organizational thinking involved in structuring those ideas coherently. The classroom implication is that AI-assisted brainstorming and planning should be an explicitly designed, teacher-structured phase of EFL writing instruction rather than an unsanctioned peripheral activity. It is precisely at this ideational threshold that the transformation from silent thinker to active writer begins: not when a student submits a finished essay, but when they first find the words to commit an internal thought to the page.

Theme 2: Vocabulary and Sentence Support

The second theme, corresponding to the highest-scoring dimension (Dimension B: M = 3.89) and present in 86.7% of responses, concerned linguistic scaffolding at the level of vocabulary and sentence construction. A consistent and educationally significant pattern emerged: students knew what they wanted to say conceptually but lacked the English words or structures to say it, turning to the chatbot as an immediate linguistic bridge:

"I already know what I mean but I don't know the exact words in English. I describe the concept to the AI and it gives me the vocabulary I need – sometimes even better words than I expected" – (R6)

"When my grammar is wrong or when my sentence sounds unnatural, I ask the AI to help me make it sound more professional" – (R11)

From Vygotsky's (1978) perspective this illustrates scaffolding at the gap between conceptual competence and linguistic performance which is the distance between what a student can think and what they can encode in the target language, which defines much of the EFL writing experience. Krashen's (1982) Affective Filter Hypothesis adds that vocabulary anxiety – fear of not knowing the right word – is one of the most immediately paralyzing forms of writing apprehension, and the chatbot's instant non-judgmental response lowers this barrier precisely when it is most obstructive. Yan and Zhang (2024) confirmed that L2 writers engaged with ChatGPT's

language suggestions critically and selectively rather than wholesale in which is a finding indicating that effective vocabulary AI use is evaluative rather than substitutive.

Theme 3: Writing Confidence and Reduced Writing Anxiety

The third theme brought together Dimensions C (M = 3.61) and D (M = 3.57), present in 80.0% and 73.3% of responses respectively. While theoretically – confidence as a stable self-perception and anxiety as a situational in-the-moment state – these constructs operated together in participants' accounts. AI chatbot availability simultaneously built confidence and reduced anxiety, consistently through the same mechanism: the non-judgmental, patient, and socially costless nature of interaction with the chatbot:

"Before I used AI chatbots I always felt like my writing was not good enough to show the teacher. Now I feel I can at least produce something I am not ashamed of" – (R15)

"The AI never makes me feel bad when my English is wrong. It just helps me fix it without making me feel stupid" – (R8)

"With a human teacher I am always worried about making mistakes in front of the class. With the AI I can make as many mistakes as I want and it doesn't matter" – (R21)

This provides direct empirical support for Krashen's (1982) Affective Filter Hypothesis: the chatbot lowers the affective filter not through reassurance but through structurally removing the social and evaluative stakes that generate writing anxiety in EFL contexts. Vygotsky's (1978) framework adds a developmental dimension: confidence built through repeated successful scaffolded interactions can gradually internalize, shifting previously AI-assisted tasks into independent capability. These findings are consistent with Almusharraf et al. (2025), who found self-efficacy among the most positively affected constructs in EFL AI writing perceptions, and with Yu (2024), who found a significant negative correlation between AI-assisted writing frequency and writing anxiety. Affectively, this is what the transition from silent thinker to active writer looks like: not a sudden gain in linguistic competence, but a gradual lowering of the internal barrier that had been keeping genuine ideas from reaching the page.

Theme 4: Enhanced Engagement and Participation

The fourth theme, corresponding to Dimension E (M = 3.54) and present in 66.7% of responses, extended the affective findings into behavioral outcomes. Students described writing more frequently, attempting more ambitious tasks, and persisting through difficulty when AI support was available by describing a meaningful shift from avoidance to engagement:

"Before AI chatbots I avoided writing tasks as much as possible. Now I actually look forward to them sometimes because I know I have a tool to help me when I need it" – (R14)

"I used to give up when I got stuck. Now I use the AI to help me get past the stuck part and then I keep writing" – (R22)

From Vygotsky's (1978) perspective, appropriate scaffolding does not merely help a learner complete a task but reshapes their relationship to the activity entirely. From Krashen's (1982) perspective, motivation is the third pillar of the affective filter alongside anxiety and confidence, and the present findings suggest AI chatbot support raises all three simultaneously, creating a mutually reinforcing affective system. Song and Song (2023) reported heightened motivation in ChatGPT-assisted EFL writing, and Duong and Chen (2025) found that purposeful AI engagement correlated with stronger writing outcomes which are the findings consistent with the behavioral shift documented here.

Theme 5: Students' Overall Perceptions of AI-Assisted Writing

The fifth theme, corresponding to Dimension F (M = 3.74) and present in all 30 responses (100%), concerned students' global evaluative perceptions of AI-assisted writing. The dominant sentiment was positive but rarely unconditional in which most participants qualified their favorable evaluations with self-aware caveats about appropriate versus inappropriate use, reflecting emerging AI literacy:

"AI chatbot is like a dictionary but more intelligent. It doesn't write my essay – I write it. But it helps me write it better than I could alone"
– (R7)

"It is useful but only if you use it correctly. If you just copy what AI says, you are not learning anything" – (R19)

This qualified positivity is itself theoretically significant: students are active metacognitive agents who reflect on and regulate their own AI use, not passive

recipients of scaffolding which is exactly what developmentally mature engagement with scaffolded support looks like. Students' broadly trustworthy perception of AI chatbots also supports Krashen's (1982) AFH: a tool perceived with suspicion would be unlikely to lower the affective filter regardless of its functional capabilities. These findings are consistent with Slamet (2024) and Escalante et al. (2023), both of whom documented broadly positive EFL student perceptions of AI tools in writing contexts.

Theme 6: Ethical Concerns and the Importance of Teacher Guidance

The sixth and most critical theme brought together two inseparable issues: the ethical risks of AI chatbot use and students' need for explicit teacher guidance in navigating them. Ethical concerns were raised by 83.3% of participants and encompassed three recurring issues: academic integrity and plagiarism, overdependence and erosion of independent writing ability, and inaccurate or contextually inappropriate AI output. Simultaneously, 90.0% of participants which is the highest prevalence of any theme in the study explicitly called for more structured teacher guidance. These two concerns belong together because teacher guidance is the pedagogical response to the ethical risks:

"I feel that I cannot write without AI anymore. Before I started using it I could at least write something. Now I feel empty when I open a blank document and there is no AI" – (R9)

"Sometimes AI gives me confident-sounding sentences that are actually wrong. If I don't check them I will submit mistakes" – (R4)

"Our teachers just say we can use AI but don't explain how to use it responsibly. I want to know the rules so I can follow them" – (R2)

"Teachers should show us the difference between using AI to help us think and using AI instead of thinking. Those are two very different things" – (R12)

These findings complicate the Vygotskian scaffolding account applied in earlier themes. Vygotsky's (1978) framework assumes scaffolding is progressively faded as independent capability grows. Without deliberate instructional design, however, AI scaffolding does not fade which it deepens into the very dependency these students describe. This is a genuine theoretical tension: the constant availability that makes AI chatbots powerful ZPD scaffolds also makes them uniquely susceptible to the failure

mode of permanent dependency. Cotton et al. (2024) identified cognitive offloading and academic misconduct as the central risks of AI integration in higher education; Amoush and Alhosban (2025) catalogued overdependence and inaccurate AI output specifically in EFL writing; and Jung et al. (2026) found these tensions structurally embedded in AI-integrated EFL writing classrooms. The teacher's role, these data make clear, is not merely to permit or prohibit AI use but to take on an active meta-scaffolding function: structuring conditions of use, designing deliberate scaffolding fading, teaching critical verification strategies, and providing the explicit ongoing guidance that the students in this study are actively requesting. Only under these conditions can the transformation from silent thinker to active writer be sustained as a genuine developmental achievement rather than an AI-dependent performance.

CONCLUSION

This study investigated how AI chatbots support idea expression among EFL university students and how those students perceive that support. The quantitative findings revealed consistently positive perceptions across all six dimensions, with a grand mean of $M = 3.71$ (High). Most highlighted that Idea Generation and Brainstorming ($M = 3.88$) and Vocabulary and Sentence Support ($M = 3.89$) emerged as the highest-scoring dimensions, confirming that AI chatbots are most effective precisely at the foundational stage where EFL writing difficulty is most acute which is the movement from unformed thought to articulated English text. Qualitatively, students described AI chatbots as brainstorming partners, organizational scaffolds, and linguistic bridges, a pattern that aligns directly with Vygotsky's (1978) concept of scaffolding within the Zone of Proximal Development. These are precisely the mechanisms through which silent thinkers – EFL learners who carry genuine ideas but cannot yet bridge the gap between internal thought and written English – are supported in becoming active writers.

Beyond cognitive support, the study documented meaningful affective benefits. Students reported reduced writing anxiety ($M = 3.57$) and increased confidence ($M = 3.61$), consistent with Krashen's (1982) Affective Filter Hypothesis: the non-judgmental nature of AI interaction structurally removes the social and evaluative pressures that inhibit EFL idea expression. Enhanced engagement ($M = 3.54$) further indicated a behavioral shift from avoidance to active participation. At the same time, 83.3% of participants acknowledged risks of overdependence and 90.0% explicitly called for

structured teacher guidance, revealing a student body that is metacognitively aware of both the benefits and the limits of AI-assisted writing. Taken together, these cognitive and affective findings confirm that when AI chatbots are used responsibly and with deliberate teacher guidance, they hold genuine potential to transform silent thinkers into active writers – not by writing for students, but by lowering the barriers that have been preventing students from writing for themselves.

SUGGESTIONS

For EFL writing instructors, the findings strongly suggest that AI chatbot use should be integrated deliberately rather than left to students' unsanctioned discretion. Teachers are encouraged to design AI-assisted brainstorming and pre-writing phases as explicit, structured components of writing instruction with clear guidelines distinguishing between using AI to support one's own thinking and using AI as a substitute for it. Scaffolding fading sequences, in which AI support is progressively reduced as independent capability grows, should be built into course design to prevent the dependency patterns documented in this study.

For future researchers, this study points toward several productive directions. Longitudinal designs that track changes in independent writing ability over sustained periods of AI chatbot use would help clarify whether the scaffolding benefits documented here genuinely transfer to autonomous performance. Observational and interaction-log-based methods that capture the real-time process of AI-assisted idea expression rather than relying solely on self-reported perceptions would provide a richer and more ecologically valid picture of how EFL learners actually engage with these tools during writing. Comparative studies across different EFL contexts, institutions, and AI platforms would also strengthen the generalizability of findings in this rapidly evolving field.

REFERENCES

- Almusharraf, A., Bailey, D., Almusharraf, N., & Alotaibi, T. (2025). Students' Perceptions of Generative AI in EFL Writing: Strategies, Self-Efficacy, Satisfaction and Behavioural Intention. *Australasian Journal of Educational Technology*, 41(5).
<https://doi.org/10.14742/ajet.10045>

- Amoush, K. H., & Alhosban, A. A. (2025). Mastering EFL Writing With ChatGPT: A Systematic Review of Benefits, Challenges, and Best Practices. *Theory and Practice in Language Studies*, 15(8). <https://doi.org/10.17507/tpls.1508.14>
- Braun, V., & Clarke, V. (2006). Using Thematic Analysis in Psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
<https://doi.org/10.1191/1478088706qp063oa>
- Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and Cheating: Ensuring Academic Integrity in the Era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228–239.
<https://doi.org/10.1080/14703297.2023.2190148>
- Duong, T. N. A., & Chen, H. L. (2025). An AI Chatbot for EFL Writing: Students' Usage Tendencies, Writing Performance, and Perceptions. *Journal of Educational Computing Research*, 63(2), 406–430.
<https://doi.org/10.1177/07356331241312363>
- Escalante, J., Pack, A., & Barrett, A. (2023). AI-Generated Feedback on Writing: Insights into Efficacy and ENL Student Preference. *International Journal of Educational Technology in Higher Education*, 20(1). <https://doi.org/10.1186/s41239-023-00425-2>
- Gao, J., Zhang, J., & Li, Y. (2025). Do AI Chatbot-Integrated Writing Tasks Influence Writing Self-Efficacy and Critical Thinking Ability? An Exploratory Study. *Computers and Education: Artificial Intelligence*, 9.
<https://doi.org/10.1016/j.caeai.2025.100472>
- Jung, J., Joo, S., Park, H., & Han, I. (2026). Challenges of Integrating ChatGPT into EFL Writing: An Activity Theory Framework. *Computer Assisted Language Learning*.
<https://doi.org/10.1080/09588221.2026.2617386>
- Kasneci, E., Sessler, K., Uchemann, S. K. ", Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Unnemann, S. G. ", Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Urgen Pfeffer, J. ", Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., ... Kasneci, G. (2023). ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education.
- Krashen, S. D. (1982). *Principles and Practice in Second Language Acquisition*. Pergamon Press.

- Mahapatra, S. (2024). Impact of ChatGPT on ESL Students' Academic Writing Skills: A Mixed Methods Intervention Study. *Smart Learning Environments*, 11(1).
<https://doi.org/10.1186/s40561-024-00295-9>
- Slamet, J. (2024). Potential of ChatGPT as a Digital Language Learning Assistant: EFL Teachers' and Students' Perceptions. *Discover Artificial Intelligence*, 4(1).
<https://doi.org/10.1007/s44163-024-00143-2>
- Song, C., & Song, Y. (2023). Enhancing Academic Writing Skills and Motivation: Assessing the Efficacy of ChatGPT in AI-Assisted Language Learning for EFL Students. *Frontiers in Psychology*, 14.
<https://doi.org/10.3389/fpsyg.2023.1260843>
- Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press.
- Wang, C. (2025). Exploring Students' Generative AI-Assisted Writing Processes: Perceptions and Experiences from Native and Nonnative English Speakers. *Technology, Knowledge and Learning*, 30(3), 1825–1846.
<https://doi.org/10.1007/s10758-024-09744-3>
- Yan, D., & Zhang, S. (2024). L2 Writer Engagement with Automated Written Corrective Feedback Provided by ChatGPT: A Mixed-Method Multiple Case Study. *Humanities and Social Sciences Communications*, 11(1).
<https://doi.org/10.1057/s41599-024-03543-y>
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic Review of Research on Artificial Intelligence Applications in Higher Education – Where are the Educators? In *International Journal of Educational Technology in Higher Education* (Vol. 16, Issue 1). Springer Netherlands.
<https://doi.org/10.1186/s41239-019-0171-0>
- Ziqian Yu. (2024). The Study of the Correlation Between AI-Assisted EFL Writing Frequency and Writing Anxiety Among Chinese University Students. *Academic Journal of Humanities & Social Sciences*, 7(7).
<https://doi.org/10.25236/ajhss.2024.070702>