



ENHANCING METACOGNITIVE SKILLS THROUGH AI-ASSISTED PROMPTING IN EAP READING AND WRITING

Omonova Umida

Tashkent Perfect University

Trainee-teacher

omonovaumida1604@gmail.com

Annotation: English for Academic Purposes (EAP) students often read and write without deliberately planning, monitoring, or evaluating their strategies, which limits the transfer of skills across disciplines. This article examines how structured prompting of generative artificial intelligence (AI) tools can strengthen learners' metacognitive skills in EAP reading and writing. Drawing on Flavell's model of metacognition, Zimmerman's cycle of self-regulated learning, and scaffolding theory, the study proposes a prompt framework in which learners use AI not as an answer generator but as a thinking partner that asks questions, models strategies, and gives feedback on process. A mixed-methods, quasi-experimental design is described for a ten-week intervention with university EAP students, combining the Metacognitive Awareness Inventory, rubric-scored reading and writing tasks, prompt logs, and reflective interviews. The article reports illustrative results showing larger gains in metacognitive awareness, reading comprehension, and writing quality in the AI-prompting group than in a comparison group, along with qualitative evidence of more frequent planning, monitoring, and revision behaviors. Examples of effective and ineffective prompts are presented in tabular form to guide classroom practice. The discussion addresses risks of over-reliance, inaccurate output, and academic integrity, and offers recommendations for task design, assessment, and teacher training. The article concludes that prompting, when explicitly taught and linked to reflection, can make invisible cognitive processes visible and so help EAP learners become more autonomous readers and writers.



Keywords: metacognition; English for Academic Purposes; generative AI; prompting; self-regulated learning; academic reading; academic writing.

Students entering English-medium universities must read dense, discipline-specific texts and produce writing that meets conventions of argument, evidence, and register. EAP courses exist to prepare them for these demands, yet many learners approach academic tasks mechanically. They read from the first line to the last, translate unfamiliar words one by one, write a single draft, and submit it without systematic revision. What is often missing is not language knowledge alone but awareness and control of the learning process itself, which is the core of metacognition (Flavell, 1979). Metacognition is widely recognized as a strong predictor of academic success and a key ingredient of self-regulated learning (Pintrich, 2002; Zimmerman, 2002). In second language contexts, learners who can name their strategies, choose among them, and evaluate their effectiveness tend to read more flexibly and write more purposefully (Anderson, 2002; Wenden, 1998). However, metacognition is difficult to teach because the processes involved are internal and largely invisible to teachers and, frequently, to students themselves. The rapid spread of generative AI tools such as large language models has created both a new instructional opportunity and a new risk. These tools can offer immediate, individualized dialogue about texts and drafts at any time of day (Kasneci et al., 2023). Yet unguided use can encourage cognitive offloading: students request a summary or a finished paragraph and skip the thinking that EAP instruction is meant to develop (Barrot, 2023; Kohnke et al., 2023). The central argument of this article is that the quality of the prompt determines whether AI replaces or stimulates metacognition. When prompts are designed so that the AI asks, models, and responds to the learner's own reasoning, the tool can serve as an external scaffold for planning, monitoring, and evaluating. Three research questions guide the study. First, does a structured AI-prompting intervention improve EAP learners'



metacognitive awareness? Second, does it improve their reading comprehension and writing quality? Third, how do learners describe and use prompting for metacognitive regulation?

2. Topicality and Theoretical Framework

2.1 Why the topic matters now

The topic is timely for several connected reasons. Universities worldwide are revising their policies on generative AI, and EAP programs are often the first place where students receive guidance on responsible use. Teachers therefore need evidence-informed approaches rather than blanket bans or uncritical enthusiasm (Kohnke et al., 2023). At the same time, if AI can produce acceptable text in seconds, assessing only the final product is no longer sufficient, and process-oriented outcomes such as the ability to plan, monitor, and revise become the most defensible learning goals (Flower & Hayes, 1981; Hyland, 2006). Research on second language reading also shows that strategy knowledge rarely becomes strategy use without explicit, repeated practice and reflection (Grabe & Stoller, 2019), and AI dialogue offers a low-anxiety space for such practice. Finally, writing an effective prompt requires learners to define a goal, specify a context, anticipate what information is needed, and judge the response. These are the same operations that constitute metacognitive control, which makes prompting a natural vehicle for developing them.

2.2 Metacognition

The study rests on three complementary theoretical foundations. The first is metacognition, a concept introduced by Flavell (1979), who defined it as people's knowledge and regulation of their own cognitive processes. Flavell distinguished metacognitive knowledge, which covers beliefs about persons, tasks, and strategies, from metacognitive experiences, such as the sudden feeling of not understanding a sentence, and from the goals and actions learners take in response. Schraw and Dennison (1994) later operationalized the construct in two broad components: knowledge of cognition, which includes declarative, procedural, and conditional knowledge, and regulation of cognition, which includes planning, information management, monitoring, debugging, and evaluation. Pintrich (2002) argued that this knowledge should be explicitly taught



and assessed, because learners who understand how tasks differ and which strategies fit them are better able to adapt in new situations. In second language learning, Wenden (1998) and Anderson (2002) showed that successful learners hold accurate beliefs about language tasks and orchestrate strategies flexibly, whereas weaker learners tend to apply a narrow set of habits regardless of the task. For EAP reading and writing, metacognition therefore explains why learners with similar language proficiency can perform very differently on the same academic task.

2.3 Self-regulated learning

The second foundation is self-regulated learning, which places metacognition within a broader cycle of motivation, behavior, and reflection. Zimmerman (2002) describes self-regulation as a cyclical process of three phases. In the forethought phase, learners analyze the task, set goals, and plan strategies. In the performance phase, they apply strategies while monitoring their progress and managing attention. In the self-reflection phase, they judge their results, attribute success or failure to causes, and adjust future approaches. Self-regulated learners are not simply more intelligent or more motivated; they are more systematic in how they use feedback to improve the next attempt. This cyclical model maps neatly onto AI-assisted prompting. A goal-setting prompt supports forethought, a comprehension-checking prompt supports performance control, and a feedback-on-process prompt supports reflection. In the framework used in this study, each stage of the cycle therefore corresponds to a distinct type of learner request, so that the sequence of prompts itself becomes a record of the learner's self-regulation.

2.4 Scaffolding and the zone of proximal development

The third foundation is sociocultural theory, in particular Vygotsky's (1978) concept of the zone of proximal development and its practical elaboration as scaffolding by Wood et al. (1976). Vygotsky proposed that learning occurs most effectively in the space between what a learner can do independently and what the learner can do with guidance from a more knowledgeable other. Wood et al. described the tutor's role as simplifying the task, keeping the learner focused on the goal, highlighting critical features, and



controlling frustration, with 4 support withdrawn as competence grows. A well-designed AI interaction can imitate several of these moves. It can ask a clarifying question, offer a hint rather than an answer, and respond contingently to what the learner has just said. However, AI does not automatically fade its support, and it does not know the learner's history. The teacher must therefore instruct the tool to hint rather than tell and must plan the gradual withdrawal of AI assistance so that the learner internalizes the questions the tool models. Taken together, the three theories explain both why prompting may develop metacognition and why it may fail when the learner's own thinking is bypassed.

3. Methodology

3.1 Design, participants, and context

The study uses a mixed-methods, quasi-experimental pretest-posttest design with a comparison group. Quantitative data measure change in metacognitive awareness and in reading and writing performance, while qualitative data from prompt logs, reflective journals, and interviews explain how learners used prompting. Quasi-experimental assignment was necessary because intact EAP classes could not be randomized. The intended sample is 60 undergraduate EAP students at upper-intermediate level (approximately CEFR B2), enrolled in two parallel sections of a ten-week course. One section ($n = 30$) forms the experimental group and the other ($n = 30$) the comparison group. Both sections share the same teacher, syllabus, and weekly contact hours, which limits instructor effects. Participation is voluntary, informed consent is obtained, grades are not affected by research participation, and all data are anonymized. The comparison group receives the AI-prompting training after the study ends.

3.2 The intervention

The experimental group received weekly training in a four-stage prompting framework aligned with the self-regulated learning cycle. In the plan stage, learners state the task goal and ask the AI to help them choose a strategy without producing content. In the monitor stage, they ask the AI to pose comprehension or coherence questions during reading or drafting and then answer those questions themselves. In the evaluate stage, they submit their own answer or draft and request feedback on



process and criteria rather than a rewrite. In the reflect stage, they ask the AI to help them identify what worked and what to change next time, and they record the conclusions in a journal. The comparison group completed identical reading and writing tasks with conventional strategy instruction and teacher feedback but without AI use. In the experimental group, AI use was limited to the framework and was monitored through exported prompt logs. Teachers also discussed the limitations of AI, including inaccurate statements and 5 invented references, and required learners to verify any factual claim.

3.3 Instruments and analysis

Metacognitive awareness was measured before and after the intervention with an adapted version of the Metacognitive Awareness Inventory (Schraw & Dennison, 1994), a 52-item instrument rated on a five-point scale. Reading comprehension was measured with a 30-item test based on a discipline-neutral academic article, combining multiple-choice, short-answer, and inference questions. Writing quality was measured with a 300 to 350-word argumentative essay scored on a 20-point analytic rubric covering task response, organization, language, and use of evidence. Two raters scored the essays independently, and agreement was checked before the scores were averaged. In addition, weekly prompt logs and reflective journals were collected from the experimental group, and ten participants took part in 20-minute semi-structured interviews after the course. Quantitative data are analyzed with analysis of covariance (ANCOVA) using pretest scores as the covariate, and effect sizes are reported as partial eta squared. Prompt logs are coded for metacognitive function (planning, monitoring, evaluating, reflecting, or non-metacognitive requests such as 'write this for me'), and interviews and journals are analyzed thematically.

4. Results

Note: The figures in this section are illustrative sample data created to show how findings would be reported. They are not drawn from a real study and must be replaced with actual data.

4.1 Quantitative findings

Table 1 summarizes performance on the three measures. After controlling for pretest scores, the experimental group scored significantly higher than the comparison group on the



posttest of metacognitive awareness ($p < .001$), with a large effect (partial eta squared = .34). The largest subscale gains were in monitoring and evaluation, the processes least likely to occur spontaneously without guidance. Both groups improved in reading and writing, which is expected after ten weeks of instruction, but the experimental group showed roughly twice the gain. The strongest writing improvements were in organization and use of evidence, the criteria most closely tied to planning and revision.

6 Table 1. Pretest and posttest scores by group ($n = 30$ per group)

Measure	Group	Pretest M (SD)	Posttest M (SD)	Gain
Metacognitive awareness (1-5)	Experimental	3.12 (0.41)	3.89 (0.38)	+0.77
	Comparison	3.15 (0.44)	3.41 (0.42)	+0.26
Reading comprehension (max 30)	Experimental	17.4 (3.6)	22.8 (3.2)	+5.4
	Comparison	17.7 (3.8)	20.1 (3.5)	+2.4
Writing quality (max 20)	Experimental	11.2 (2.3)	14.9 (2.1)	+3.7
	Comparison	11.4 (2.5)	13.0 (2.3)	+1.6

4.2 Prompt-log findings Analysis of 842 prompts from the experimental group showed a clear change in how learners used the tool. During the first three weeks, 53 percent of prompts were non-metacognitive requests, such as asking the AI to translate or rewrite text. By weeks eight to ten this share had fallen to 3 percent. Over the same period, monitoring prompts rose from 12 to 31 percent, planning prompts from 14 to 27 percent, evaluating prompts from 18 to 29 percent, and reflecting prompts from 3 to 10 percent. This pattern suggests a shift from product-seeking to process-seeking use and supports the claim that prompting behavior can be taught.

4.3 Qualitative findings Thematic analysis of journals and interviews produced four themes. The first was making thinking visible: learners reported that writing a prompt forced them to state what they did not understand, and one participant noted that explaining the problem to the AI showed that the weakness was the topic sentence, not the grammar. The second theme was safe rehearsal, since learners felt less anxious asking repeated questions of a tool than of a teacher or peer. The third was critical evaluation of AI: after instruction, learners began verifying claims and noticing errors in AI replies, which they said increased their sense of responsibility



for their own work. The fourth theme was the temptation to offload. Several learners admitted that under time pressure they were tempted to ask the AI to produce the answer, which confirms the need for explicit norms and teacher oversight. 5. Examples of Metacognitive Prompts The examples in Table 2 show how the framework can be applied in EAP reading and writing. The guiding principle is that the learner supplies the thinking and the AI responds to it. Each metacognitive prompt names the learner's goal, includes the learner's own attempt where possible, and tells the AI what it must not do. 7 Table 2. Weak versus metacognitively oriented prompts

Stage and task	Weak prompt (offloading)	Metacognitive prompt (scaffolding)
Plan: reading	"Summarize this article for me."	"I have 30 minutes to read an 8-page article on urban planning for a seminar. Ask me three questions about my purpose, then suggest a reading strategy and explain why it fits."
Monitor: reading	"What does paragraph 4 mean?"	"Here is my two-sentence summary of paragraph 4. Do not correct it yet. Ask me two questions that test whether I understood the author's claim."
Evaluate: reading	"Is this author right?"	"I think the evidence is weak because the sample is small. Challenge my judgment with one counterargument and ask me to respond."
Plan: writing	"Write an outline for my essay on remote work."	"My thesis is that remote work reduces productivity in creative teams, and I plan three body paragraphs. Ask questions that expose gaps in my plan, but do not write the outline."
Evaluate: writing	"Make my essay better."	"Using these criteria (task response, organization, evidence), give feedback on organization only and explain your reasoning in two sentences. I will revise it myself."
Reflect	"Give me tips for writing."	"Here are my last two essays and my teacher's comments. Help me identify the one habit that caused most of my problems and suggest a goal for the next essay."

A short exchange illustrates the monitoring stage in practice. A learner who had read a section on climate migration wrote a one-sentence summary stating that people move because of floods and asked the AI to test the summary. The AI replied that the author also mentions gradual factors and



asked which the author treats as the main driver and what evidence supports it. The learner re-read the section, noticed that drought and loss of income were emphasized in the third paragraph, and revised the summary. When the AI then asked what strategy would prevent the same oversight, the learner decided to underline the main claim of each section before summarizing. In this single exchange the learner detected a comprehension gap, repaired it, and formed a plan for future tasks, which are the monitoring, debugging, and planning processes described by Schraw and Dennison (1994).

6. Discussion The illustrative findings align with theory in three ways. First, the gains in monitoring and evaluation are consistent with the view that metacognitive regulation improves when learners externalize their reasoning and receive timely feedback on it (Schraw & Dennison, 1994; Zimmerman, 2002). Second, the shift in prompt types over time shows scaffolding in action: as learners internalize the questions the AI models, their dependence on external prompting should decline (Wood et al., 1976). Third, improvements in organization and evidence use mirror classic process models of writing in which planning and revision are central (Flower & Hayes, 1981). The results should nonetheless be interpreted with caution. The effect depends on instruction: without explicit training and a requirement to reflect, many learners default to product-oriented requests, as the early prompt logs show. Generative AI also makes errors, may invent sources, and may reflect biases in its training data, so critical evaluation of output must itself be taught as a metacognitive skill (Barrot, 2023). Several pedagogical implications follow. Task design should require learners to submit their own attempt before consulting AI, which keeps the cognitive effort with the learner. Teachers should model the Plan, Monitor, Evaluate, and Reflect framework with sentence frames and gradually reduce AI use so that strategies become independent. Assessment should include process evidence, such as prompt logs, reflections, and draft histories, alongside the final product. Syllabi should define acceptable AI use and require disclosure, and activities in which students verify AI claims and sources should be



built into the course. Finally, teachers need professional development in prompt design and log interpretation so that they can guide rather than merely police AI use. The study has limitations. A quasi-experimental design cannot rule out all differences between classes, ten weeks is too short to show whether gains persist, and self-report instruments may overstate awareness. Future research should add think-aloud protocols and delayed posttests and should examine learners at different proficiency levels and in discipline-specific EAP contexts. 7. Conclusion Generative AI will not, by itself, make students more reflective readers and writers, and used carelessly it may do the opposite. This article has argued that structured prompting, grounded in theories of metacognition, self-regulated learning, and scaffolding, can turn AI into a mirror for learners' thinking. By planning with the AI, monitoring comprehension through its questions, evaluating drafts against criteria, and reflecting on their strategies, EAP students make their cognitive processes visible and therefore improvable. The framework and prompt examples presented here give teachers a practical starting point. The central pedagogical message is simple: in AI-supported EAP classrooms, the most valuable skill to teach is not how to get answers from the tool but how to ask questions that make learners think.

References

Al-khresheh, M. H. (2024). The role of presentation-based activities in enhancing speaking proficiency among Saudi EFL students: A quasi-experimental study. *Acta Psychologica*, 243, 104159.

<https://doi.org/10.1016/j.actpsy.2024.104159>

Asrifan, A., Rosmayanti, V., Vargheese, K. J., & Vivekanantharasa, R. (2025). AI-supported video projects in ESP classrooms: A pedagogical innovation for business communication. *Journal of English Language Teaching and Islamic Studies (JELTIS)*.



Astawa, P., Artini, L., & Nitiasih, P. (2017). Project-based learning activities and EFL students' productive skills in English. *Journal of Language Teaching and Research*, 8(6), 1147–1155. <https://doi.org/10.17507/jltr.0806.16>

Aysah, F., & Hamsia, W. (2025). AI-based project-based learning to improve speaking skills and confidence in ECC. *Journal of English Education and Linguistics*, 13(2). <https://doi.org/10.32682/jecell.v13i2.149>

Bakar, N. I. A., Noordin, N., & Razali, A. B. (2019). Improving oral communicative competence in English using project-based learning activities. *English Language Teaching*, 12(4), 73–84. <https://doi.org/10.5539/elt.v12n4p73>

Beckett, G. H., & Slater, T. (2019). *Global perspectives on project-based language learning, teaching, and assessment: Key approaches, technology tools, and frameworks*. Routledge.

Bell, S. (2010). Project-based learning for the 21st century: Skills for the future. *The Clearing House*, 83(2), 39–43. <https://doi.org/10.1080/00098650903505415>

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1–47. <https://doi.org/10.1093/applin/1.1.1>

Edwards, B. I., & Yahaya, W. A. J. W. (2025). Enhancing English communicative competence through AI tools in vocational higher education: A study among English for Professional Purposes (EPP) students. *Journal of Technical Education and Training*, 17(2), 137–150. <https://doi.org/10.30880/jtet.2025.17.02.010>

Hymes, D. H. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics: Selected readings* (pp. 269–293). Penguin.