Discover Computing

RESEARCH

Open Access

Interactive reading patterns and metacognitive strategies from surface to deep levels of L2 learners in GenAI-assisted critical reading

Haoming Lin[1], Wei Wei[1*], Bin Zou[2] and Ying Zheng[3]

*Correspondence:
Wei Wei
weiweitesting@hotmail.com
[1]Faculty of Applied Sciences, Macao Polytechnic University, Avenida Padre Tomás Pereira, Macao, Macao
[2]Department of Applied Linguistics, Xi'an Jiaotong-Liverpool University, Suzhou, China
[3]School of Humanities, University of Southampton, Southampton, UK

Abstract
Engaging critically with academic literature is challenging yet indispensable for research-oriented postgraduate students. Although GenAI-powered tools have been developed to summarize, evaluate, and revise academic texts, there remains a paucity of research on how these tools specifically support the reading processes of second language (L2) readers. To address this gap, this study recruited 53 postgraduate students who were tasked with drafting and revising a critical review of a published academic article using three resources: (1) critical reviews generated by a GenAI-powered tool developed on the Coze platform, (2) communications with an embedded GenAI chatbot, and (3) the original academic article. Lag Sequential Analysis (LSA) revealed three distinct patterns of learner interaction: (1) predominantly relying on GenAI-generated critical reviews with minimal rereading of the original academic article; (2) primarily engaging in careful reading of GenAI-generated critical reviews, supplemented by frequent rereading of the original academic article; and (3) extensively integrating both GenAI-generated critical reviews and direct chatbot interactions. Qualitative interviews further indicated that these interaction patterns were influenced by learners' deployment of metacognitive strategies adapted to task complexity and cognitive demands. Specifically, Pattern 1 was characterized by surface-level strategies for less challenging tasks, whereas Patterns 2 and 3 involved deeper, more extensive strategies: exploring, evaluating, and synthesizing ideas across GenAI-generated critical reviews, chatbot interactions, and the original academic article. These findings suggest that instructors should scaffold students away from sole reliance on GenAI-generated content toward multi-source evaluation to foster deeper metacognitive strategies and enhance reading depth.

Keywords AI, Human-AI interaction, Metacognitive strategies, Critical reading, Academic reading

1 Introduction

Generative Artificial Intelligence (GenAI) has emerged as a transformative tool in language education, offering affordances such as personalized feedback and interactive discussions based on learners' needs [1, 2]. Beyond general learning, GenAI-powered reading tools are capable of summarizing complex arguments [3], evaluating reading proficiency [4], and propose revisions to draft papers in an academically manner [5]. These capabilities enhance learners' engagement with challenging texts and open new avenues for developing critical reading, which involves adopting a critical perspective towards a written text and represents the application of critical thinking in reading [6–8]. In higher education, academic reading increasingly occurs in complex, multimodal, multi-document digital environments [9] featuring hyperlinked texts, interactive glossaries, and cross-referenced databases. While these affordances can deepen understanding, many L2 learners struggle with the cognitive demands of navigating non-linear structures, integrating textual and visual information, evaluating multiple sources, and maintaining focus amid peripheral content [10]. These challenges often hinder the transition from surface-level reading to higher-order comprehension, which is especially important for discerning valid content in the AI era [9, 11]. To cope with different levels of reading comprehension under varying task complexity and cognitive demand, students may deploy metacognitive strategies at different levels, shifting between more surface-oriented regulation and deeper monitoring and evaluation [12]. Accordingly, metacognitive reading strategies become crucial for sustaining attention and engaging strategically with digital texts [13, 14], as evidenced by extensive research showing that readers who effectively use these strategies achieve higher reading comprehension and that such strategies can compensate for L2 readers' insufficient proficiency [15–17].

Metacognitive reading strategies involve a cycle of sub-strategies such as planning, monitoring, and evaluation [16, 18, 19]. Although GenAI-enhanced writing tools have been shown to scaffold these sub-strategies [20], there remains limited insight into how L2 learners dynamically deploy metacognitive strategies at different levels during GenAI-assisted critical reading. Moreover, research on GenAI-assisted reading has mainly focused on end-of-task outcomes (e.g., motivation or interest) and overlooked learners' in-process behaviors and decision-making during academic reading across multiple sources [21]. To address this gap, this study explored learners' reading behavioral patterns in a GenAI-assisted context and how adaptively deployed metacognitive strategies at varying levels shaped these patterns to meet specific cognitive demands.

2 Literature review

2.1 Metacognitive reading strategies and strategy depth

Metacognitive reading strategies refer to the conscious methods readers use to plan, monitor, and evaluate their comprehension during reading [18, 19]. While previous research has addressed these sub types of metacognitive strategies, it is equally important to consider the depth of these strategies in understanding metacognitive processes [22]. Previous scholars, such as Huang, et al. [23], have emphasized the need for flexible adaptation of metacognitive strategies according to cognitive demand and task difficulty, which can be categorized as easy versus difficult based on Yang and Bai's [24] classification. In this context, Stanton, et al. [12] distinguished surface metacognitive strategies as those involving recalling, copying, or paraphrasing content, from deep strategies, which involve connecting ideas, generating novel solutions, and applying skills in innovative ways. Similarly, Ku and Ho [25] characterized surface strategies as merely identifying problems or paraphrasing information, while deep strategies include proposing solutions and advancing one's own understanding.

2.2 Related work on AI-assisted reading and metacognitive strategies

After reviewing the existing literature, emerging evidence suggested the potential effectiveness of different AI tools for supporting different levels of metacognitive sub-strategies in reading. Surface-level planning strategies involve noting procedural requirements or basic goals of a task, while deep-level planning strategies entail articulating explicit, actionable, goal-directed plans [25]. Existing studies suggested that AI tools enable users to apply planning strategies at different levels within a more scaffolded environment. For instance, CReBot posed critical questions that prompted advanced (but not novice) readers to illustrate their reading goals, a form of surface-level planning strategy, which contributed to enhanced reading engagement compared to a static paper-based guideline [26]. Yang, et al. [27] integrated a BERT-driven test item generation system that allowed students to preview test items and freely arrange their answering order, representing a deep-level planning strategy involving flexible goal-setting and organization of plans. Students who repeatedly used the system improved their reading performance.

Regarding monitoring strategies, the surface level involves simply detecting comprehension gaps or changes, while the deep level entails actively tracking understanding by integrating and connecting ideas [25, 28]. Current literature has revealed that AI-powered reading tools offer mixed results on whether they support or distract from these strategies of varying depth. For example, a chatbot presenting counterarguments and rebuttals about genetically modified organisms (GMOs) improved users' attitudes towards GMOs, although it performed no better than paper-based materials in supporting surface-level monitoring strategies, such as self-monitoring of attitude changes [29]. In contrast, Liu, et al. [30] found that an AI tool helping learners compare their questions with AI-generated personalized questions improved real-time knowledge acquisition and fostered deep-level monitoring strategies, including comprehension monitoring through active reflection and integration of information, more than traditional methods. However, Toyokawa, et al. [31] observed that timing functions and visual effects in the AI-supported reading system might hinder participants' application of both surface- and deep-level monitoring strategies to track their concentration levels.

Surface-level evaluation strategies involve acknowledging problems without taking further action, whereas deep-level evaluation strategies are associated with assessments that lead to revising one's thinking in order to generate novel interpretations or solutions [12, 25]. Relevant research on AI-powered reading tools has demonstrated improvements in students' evaluation strategies; however, the depth of strategy use tends to vary according to students' performance levels. Cheung, et al. [11] conducted a six-week ChatGPT intervention which raised secondary students' ability to evaluate scientific text credibility, though few demonstrated higher-level evaluation skills. In another study, Nguyen, et al. [32] tested doctoral students in an AI-assisted reading and writing task and indicated that low-performing students tended to copy content without deep evaluation, while high-performing students critically assessed and integrated AI-generated text.

Recent studies have begun to investigate how GenAI-supported systems may enhance learners' use of metacognitive sub-strategies during reading more effectively than traditional instructional settings [1, 33]. While these studies mark an important step forward, the literature still lacks a systematic examination of the continuum from surface-level to deep-level metacognitive strategies within AI-enhanced learning environments. In particular, little attention has been paid to how specific sub-strategies progress along this continuum. For example, Pan, et al. [21] briefly acknowledged the distinction between surface and deep strategies but did not explore it in depth. Moreover, most research has prioritized outcome-based metrics (e.g., reading comprehension scores) over process-oriented analyses that capture how learners actually read academic texts in real time. This gap highlights the need for further research on how learners behave in GenAI-assisted reading contexts and how they develop and shift from surface-level to deep-level metacognitive strategies.

2.3 Research questions

A review of the literature has revealed three key gaps. First, little is known about the metacognitive strategies learners deploy while using GenAI during academic reading; without such knowledge, it is difficult to adapt the technology to foster critical reading skills or to give learners clear, ethically grounded guidance. Second, although AI's benefits for basic comprehension are well documented, how L2 learners interact with GenAI-powered reading tools to support critical reading, an activity that demands higher-order thinking and sophisticated metacognitive strategy use, remains largely untested [14]. The few available studies [11, 26] compare AI and non-AI conditions but rarely probe learners' cognitive processes. Third, identifying effective strategies and key factors when using GenAI is essential for developing L2 readers' AI literacy and for informing the design of intuitive, supportive tools. These gaps motivated the present study and led to the following two research questions.

RQ1: How do L2 learners interact with the GenAI-powered reading tools during the academic critical reading process?

RQ2: Why do they interact with the GenAI-powered reading tools in the identified way?

3 Research design

3.1 Participants and research context

This study was conducted at a leading research-intensive university in eastern China. Fifty-three participants, aged 22.25 to 27.67 years, were recruited through snowball sampling, with 22 initially contacted via three university teachers and then these 22 invited 31 individuals to join the study. The sample consisted of 66% first-year, 23% second-year, and 11% third-year master's students, with 87% female and 13% male. All participants were pursuing master's degrees in English education or applied linguistics, with language proficiency at Levels 7–8 on China's Standards of English Language Ability, corresponding to B2–C1 on the Common European Framework of Reference for Languages. Prior to the study, all participants provided informed consent and reported one month to one year's experience using GenAI tools, such as DeepSeek, Ernie Bot, or Kimi chatbot, to support their learning. Additionally, they had 4 to 6 years of training in academic reading and critical thinking skills.

3.2 GenAI-powered reading tool

In the present study, we developed a GenAI-powered tool to support personalized academic critical reading and interactive discussion. This tool was built on Coze, an AI application development platform by the ByteDance company, and integrated the large

Fig. 1 The document upload interface

Table 1 Wallace and Wray's (2021) critical reading framework

No.	Aspects of critical reading
1	Purposes and research design
2	The contributions of the article
3	Main arguments
4	The degree of generalizability of the findings
5	The adequacy of the evidence to support the main arguments
6	The effectiveness of linking theoretical concepts with their main arguments
7	Moral and value stance
8	The degree to which arguments are substantiated or contested by existing research
9	The degree of alignment of arguments with the reader's experience
10	The comprehensive evaluation of the article

language models of DeepSeek R1. The tool comprises three main modules: (1) It allows each student to freely upload the selected academic article (see Fig. 1 for the upload interface). After uploading, the tool uses the LinkReaderPlugin to extract plain text from the article and simultaneously invokes ten DeepSeek R1 models, each analyzing a specific aspect of critical reading based on Wallace and Wray's critical reading framework [34] (see Table 1), using the article content as input. The results are then linked and displayed in the subsequent interface. (2) The user interface consists of ten interactive buttons, each corresponding to one aspect of critical reading. When a button is clicked, the relevant GenAI-generated critical review is displayed (see Fig. 2). (3) A DeepSeek-powered chatbot is embedded on the right side of the interface (see Fig. 2), providing students with opportunities to engage in discussions about the article content. This chatbot is pre-configured to analyze user-uploaded academic documents and includes a memory function that retains and utilizes the conversation history for enhanced interaction.

3.3 Research process

The research process consisted of four stages over a month (see Fig. 3).

In the first stage, pre-training (lasting 120 min), the goals and procedures were introduced. Subsequently, the researcher instructed students in searching for academic papers, downloading full texts, and writing a critical review based on reading an

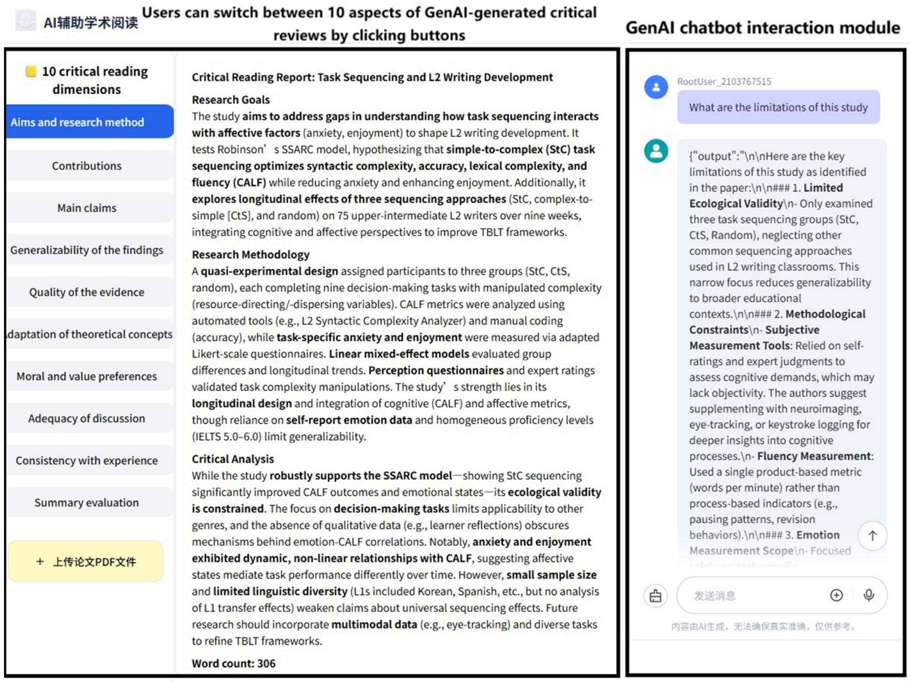

Fig. 2 The main user interface

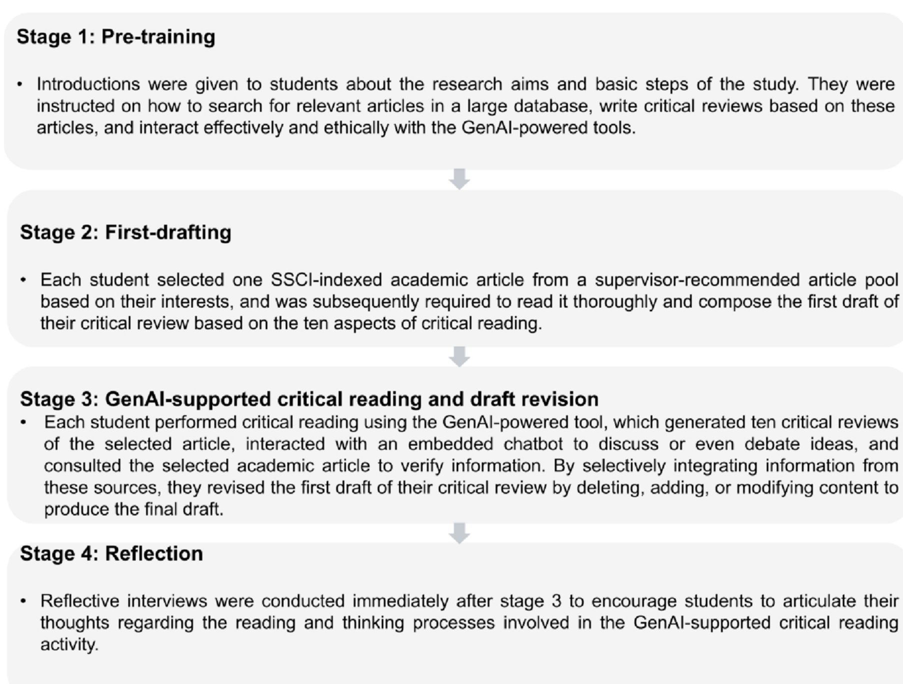

Fig. 3 The research process

academic paper. Although students reported prior experience with GenAI chatbots, they were still guided to use the GenAI-powered tool, with explicit instruction on academic integrity and responsible AI use including avoiding plagiarism.

In the second stage, first-drafting, students had two weeks to complete two tasks: (1) selecting an SSCI-indexed empirical article of interest from a supervisor-recommended article pool that aligned with their research interests to simulate an authentic academic context, and (2) critically reading the selected article and writing a critical review following Wallace and Wray's framework [34]. Specifically, at this stage, the participants completing the reading and first-draft writing in a teacher-supervised computer lab. In the third stage, GenAI-supported critical reading and draft revision (three weeks), students engaged in multiple-source critical reading and revised their drafts in a computer lab. They were given access to a GenAI-powered tool that included a module for automatically generating ten critical reviews of the selected article and an embedded GenAI chatbot, as well as the digital file of the selected academic article, to support revision. During this period, students were free to decide whether and how to use these resources and to allocate their time to reading and how to revise (e.g., paraphrase in their own words or propose contrasting ideas) until they were satisfied with their work, and both their initial draft and revised version were submitted to their supervisors. Finally, in the fourth stage, reflection, immediately after the performance stage, interviews were conducted to encourage students to reflect on their metacognitive processes. The second to final stages of the procedure were designed based on Zimmerman's three-phase model [35], which emphasizes metacognitive processes: forethought, performance, and self-reflection. Specifically, the second stage, first-drafting, corresponded to the forethought phase, as students planned their reading and reviewing strategies; the third stage, GenAI-supported critical reading and draft revision, aligned with the performance phase, where students monitored and adapted their thinking during reading and revision; and the fourth stage, reflection, corresponded to the self-reflection phase, during which students evaluated and reflected on their metacognitive processes.

3.4 Data collection

Two sources of data were collected: (1) video recordings of students' operations during the experiment and (2) audio recordings of students' reflections from interviews. A total of 53 videos were recorded using OBS Studio, a screen-capture software on Windows, to capture students' actions. These videos averaged 1.56 h in duration (SD = 0.50, Min = 1.00, Max = 3.34). Reflective interviews were conducted immediately after the experiment, allowing students to describe metacognitive strategies that explained their behavioral patterns when interacting with GenAI-powered tools. During the interviews, students reviewed their self-written critical reviews, GenAI-generated critical reviews, selected academic articles, chatbot logs, and checked their screen-recorded videos to recall their reading and thinking processes. The following questions, adapted from Yao, et al. [20] and adjusted for the reading context, were posed to elicit reflections: (1) What were your reading purpose and plans before integrating GenAI tools? (2) While using these tools, did you track your reading process? What were you thinking? (3) Did you notice any mistakes in your self-written or GenAI-generated critical reviews? (4) What is your perception of using GenAI-powered tools, and what do you see as their possible advantages and disadvantages?

3.5 Data analysis

Firstly, to address RQ1, lag sequential analysis (LSA) was performed, and we followed three phases of data coding and analysis. In the first phase, 82.68 h of video-recordings of students' operations were examined by the first and second authors to establish a coding framework (see Table 2), which included four types of learning behaviors: rereading the self-selected academic article (Reread-Article), communicating with the GenAI chatbot (GenAI-bot1 to GenAI-bot10), careful reading GenAI-generated critical reviews across ten aspects (CarefulRead-GenAI-R1 to CarefulRead-GenAI-R10, each representing a specific aspect), and rewriting self-written critical reviews on these ten aspects (Rewrite-SWCR-1 to Rewrite-SWCR-10). In the second phase, the screen-capture video-recordings were imported into the OKC software developed by Zhang et al. [36], which supports continuous playback, keyboard shortcuts for marking the onset and offset of predefined behaviors, and text annotation of each behavioral code. The first author then watched all recordings and coded students' behaviors according to the established coding framework (see Table 2). Operationally, rereading the self-selected academic article (Reread-Article) was coded when the mouse cursor remained on the window displaying the academic article and the screen-recording showed reading-related actions (e.g., scrolling, page navigation, locating sections via mouse); communicating with the GenAI chatbot (GenAI-bot1 to GenAI-bot10) when the cursor moved within or clicked on the chatbot interface; careful reading of GenAI-generated critical reviews (CarefulRead-GenAI-R1 to CarefulRead-GenAI-R10) when the cursor stayed within the interface showing the GenAI-generated reviews on a given aspect and participants showed reading actions (e.g., scrolling or text selection); and rewriting self-written critical reviews (Rewrite-SWCR-1 to Rewrite-SWCR-10) when the cursor was positioned within the student's own critical review document and visible text entry occurred. Following Van Braak et al. [37], behavioral changes lasting less than 2 s were not coded as distinct events; the previous code was retained to avoid over-fragmentation. After coding, OKC automatically exported an event log listing, in the original playback order, each behavioral code together with its start time, end time, and duration, thus providing the sequential data required for LSA. The second author then independently reviewed the OKC files, comparing the recordings with the coded events, and any disagreements were resolved through discussion. In the final phase, behavioral data, including frequencies and sequences, were input into GSEQ 5.1 to calculate z-scores for transitional sequence significance; transitions with Z-score > 1.96 were retained to plot the state-transition diagram, and the three behavioral patterns were derived by grouping these significant transitions into recurring pathways, following the procedure reported by Chen and Chang [38].

Secondly, to answer RQ2, a qualitative thematic analysis was conducted to explore why students developed their interaction patterns with these tools, through the analysis of interview transcripts with a total duration of 54.25 h of recordings. This process included three steps. First, considering the integration of GenAI-powered tools to support reading, a coding framework was developed based on literature on metacognitive strategies, including planning, monitoring, and evaluation [16, 18, 19], and further incorporated the distinction between surface-level and deep-level strategy use to reflect the degree

Table 2 The coding framework of behavior

Coding	Description of behavior
Reread-Article	Rereading the original academic article
GenAI-bot1- GenAI-bot10	Communicating with the GenAI chatbot
CarefulRead-GenAI-R1 to CarefulRead-GenAI-R10	Careful reading the GenAI-generated critical reviews
Rewrite-SWCR-1 to Rewrite-SWCR-10	Rewriting the self-written critical reviews

of metacognition [24, 25]. First, considering the integration of GenAI-powered tools to support reading, we developed a coding framework based on metacognitive strategies, including planning, monitoring, and evaluation based on literature [16, 18, 19], and operationalized strategy depth [12, 25] in AI-assisted reading by linking metacognitive regulation to different levels of task complexity and cognitive demand [24] as follows: surface-level strategy use was coded when students' planning/monitoring/evaluation, typically sufficient for easy tasks and low cognitive demand, focused on restating information and identifying comprehension problems, whereas deep-level strategy use was coded when their planning/monitoring/evaluation, mobilized in response to difficult tasks and high cognitive demand, focused on synthesizing information across sources, generating new interpretations/outputs, and solving or evaluating solutions to advance understanding. Building on Shang et al. [39], critical reading aspects that mainly involve direct interpretation from the academic text (see Table 1; e.g., Wallace and Wray's critical reading aspects 2, 3, 4, 8, and 9) represent a lower level of critical reading with easier tasks and lower cognitive demand, whereas aspects requiring analysis and inference (see Table 1; e.g., Aspects 1, 5, 6, 7, and 10) involve more difficult tasks and higher cognitive demand. Second, two authors thoroughly reviewed the transcripts and identified subthemes related to both the types and levels of metacognitive strategies that shaped students' reading and interaction behaviors. Third, the first author re-read the transcripts and coded relevant expressions, while the second author independently checked the coding results. A face-to-face meeting was held to resolve any discrepancies, and the interrater agreement rate reached 88.5%. To enhance the credibility of the qualitative findings and minimize researcher bias, each interviewee was also invited to check the accuracy of our interpretation of their individual qualitative data [40].

4 Findings

4.1 RQ1: How do L2 learners interact with the GenAI-powered reading tools during the academic critical reading process?

After implementing LSA, we identified three distinctive behavioral patterns, involving careful reading GenAI-generated critical reviews, rewriting self-written critical reviews, rereading the self-selected academic article, and communicating with the GenAI chatbot, which are presented in four representative transition cycles (see Fig. 4).

Pattern 1 included a directional sequence from the academic article to self-written critical reviews and interactions between GenAI-generated and self-written critical reviews. Transition cycles Reread-Article→Rewrite-SWCR-2 and Rewrite-SWCR-2→CarefulRead-GenAI-R2→Rewrite-SWCR-2 were identified. A unidirectional path from Reread-Article to Rewrite-SWCR-2 (Z-score: 4.31) indicated students first engaged with the academic article and then accessed self-written critical reviews on the aspect of the contributions (aspect 2). Bidirectional paths from Rewrite-SWCR-2 to CarefulRead-GenAI-R2 and from CarefulRead-GenAI-R2 to Rewrite-SWCR-2, with significant Z-scores of 35.59 and 40.07, respectively, demonstrated repetitive interactions with GenAI-generated and self-written reviews. These findings implied that students involved in this pattern in a rather superficial manner, in the sense that they focused on coverage and language-level rewriting of their own reviews while rarely revisiting the article's argumentation to validate, challenge, or synthesize ideas, and treated the GenAI-generated critical review primarily as a supplementary

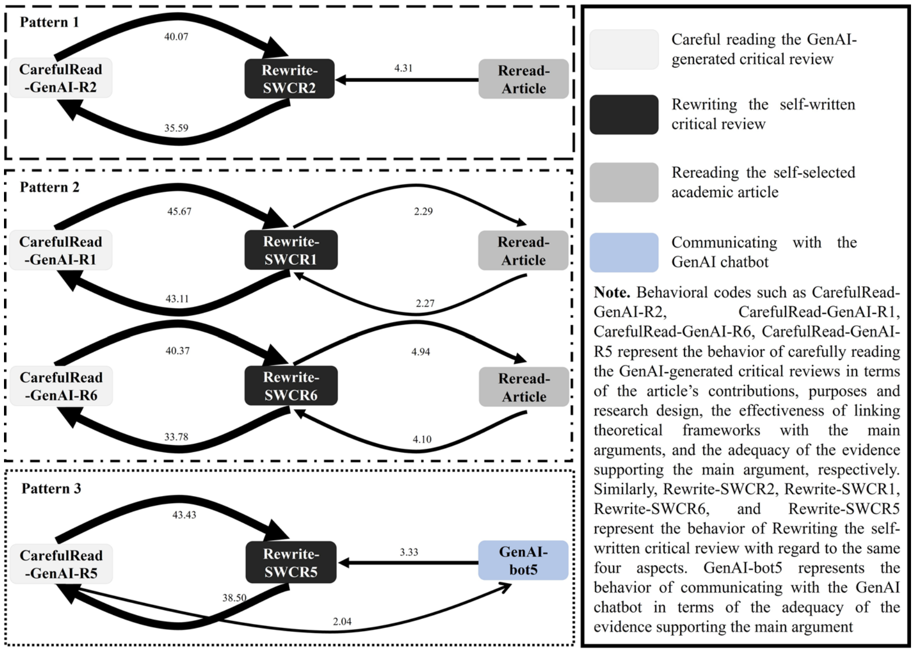

Fig. 4 The three interactive reading patterns

checklist. This indicated that their focus remained on completeness rather than on deeper synthesis.

Pattern 2 involves two bidirectional relationships: one between rewriting self-written critical reviews and careful reading GenAI-generated critical reviews, and another between rewriting self-written reviews and rereading the academic article, distinguishing it from Pattern 1. Interactive cycles like Rewrite-SWCR-1→CarefulRead-GenAI-R1→Rewrite-SWCR-1 and Rewrite-SWCR-1→Reread-Article→Rewrite-SWCR-1 were identified. Transitions between Rewrite-SWCR-1 and CarefulRead-GenAI-R1 (Z-score = 43.11) and between CarefulRead-GenAI-R1 and Rewrite-SWCR-1 (Z-score = 45.67) indicated frequent switching between self-written and GenAI-generated critical reviews, focusing on purposes and research design (aspect 1). Similarly, transition paths between Rewrite-SWCR-1 and Reread-Article (Z-score = 2.29) and between Reread-Article and Rewrite-SWCR-1 (Z-score = 2.27) revealed iterative engagement with self-written reviews and the academic article. Pattern 2 also encompassed cycles like Rewrite-SWCR-6→CarefulRead-GenAI-R6→Rewrite-SWCR-6 and Rewrite-SWCR-6→Reread-Article→Rewrite-SWCR-6, and the Z-scores from Rewrite-SWCR-6 to CarefulRead-GenAI-R6 and CarefulRead-GenAI-R6 to Rewrite-SWCR-6 (33.78 and 40.37) indicated frequent transitioning between GenAI-generated and self-written reviews on the effectiveness of linking theoretical frameworks with main arguments (aspect 6). Bidirectional sequences from Rewrite-SWCR-6 to Reread-Article (Z-score = 4.94) and Reread-Article to Rewrite-SWCR-6 (Z-score = 4.10) showed similar interactions between their reviews and the academic article. This suggested that students rewrote their reviews by comparing their critical views with intensive readings of the GenAI-generated ones and by frequently rereading the academic article for validation. The frequent shifts can be interpreted as a sustained, iterative effort to refine clarity, coherence, and argumentation through repeated cross-checking and adjustment.

Table 3 Interviewees' demographic information

Pseudonym	Gender	Age (Years)	Year of Study	Major
Cai	Female	24.53	Year 2	English Education (EE)
Huang	Female	22.94	Year 1	Applied Linguistics (AL)
Li	Female	25.45	Year 2	AL
Xiao	Female	27.53	Year 3	EE
Mao	Male	25.95	Year 1	AL
Liang	Female	24.86	Year 3	EE
Zhang	Male	24.12	Year 1	EE
Liao	Male	24.50	Year 2	AL

Pattern 3 was the only behavioral pattern that involved deeper interaction between the GenAI chatbot and the self-written critical review. A series of statistically significant transition cycles, such as CarefulRead-GenAI-R5→GenAI-bot5, GenAI-bot5→Rewrite-SWCR-5, and Rewrite-SWCR-5→CarefulRead-GenAI-R5→Rewrite-SWCR-5 were discovered. The transition CarefulRead-GenAI-R5→GenAI-bot5 ($Z = 2.04$) indicated that after careful reading the GenAI-generated critical review on evidence adequacy (aspect 5), learners would interact with the chatbot with follow-up questions. The sequence GenAI-bot5→Rewrite-SWCR-5 ($Z = 3.33$) showed that after interacting with the chatbot, the subsequent behavior involved rewriting self-written reviews on evidence adequacy. Bidirectional transitions such as Rewrite-SWCR-5→CarefulRead-GenAI-R5 ($Z = 38.50$) and CarefulRead-GenAI-R5→Rewrite-SWCR-5 ($Z = 43.43$) revealed iterative engagement with GenAI-generated and self-written reviews. The results suggested that students noticed uncertainty in the GenAI-generated critical reviews and then engaged in repeated, targeted interactions with the chatbot, such as asking for clarification, challenging the tool's interpretations, and requesting deeper explanations, to gain a more tailored understanding of the adequacy of the evidence in the selected study. This tendency to avoid returning to the academic article may indicate a reliance on dialogue-based sense-making over direct text-based verification, reflecting the perceived difficulty or effort involved in judging evidential adequacy from the article alone.

4.2 RQ2: Why do they interact with the GenAI-powered reading tools in the identified way?

The ways in which L2 learners interacted with GenAI-powered reading tools can be understood through their adaptive use of metacognitive strategies in response to the nature of the cognitive demands of different aspects of critical reading with varying task difficulty (Interviewees' demographic information is provided in Table 3).

First, in the formation of Pattern 1, learners generally perceived the critical reading aspect of contributions (aspect 2) as the least cognitively demanding, and they tended to apply metacognitive reading strategies in a relatively surface-level manner, focusing primarily on recalling key points from the academic article, paraphrasing GenAI-generated content, and broadening the range of perspectives in their own critical reviews, as demonstrated in Cai's quote. Surface-level planning strategies were evident in how learners recalled key contributions in an easy manner without extensive re-engagement with the academic article, suggesting a pre-established mental framework guiding their reading process, as shown in Cai's quote. Surface-level monitoring strategies emerged as learners compared their self-written reviews with GenAI-generated critical reviews, incorporating additional insights and refining their writing for completeness, as Cai explained. Meanwhile, learners frequently reported using surface-level evaluation strategies, ranging from checking completeness to checking grammatical errors, reflecting a shift in focus from content to form when substantive revisions were deemed unnecessary, as Cai and Huang noted.

Before re-reading, I remembered the contributions in the academic article quite clearly, so recalling its key contributions should be very easy. While reading the GenAI-generated critical review, I noticed some points I hadn't thought of before, and they made a lot of sense… After reading and revising, I felt I had finished a rather complete review (Cai).

I felt like I had already revised enough content here, so I mainly evaluated whether I made grammar mistakes in my own review. For example, I realized that one of my sentences had a subject-verb disagreement (Huang).

Second, in Pattern 2, learners perceived critical reading aspects, such as purposes and research design (aspect 1), and the effectiveness of linking theoretical frameworks with main arguments (aspect 6), as slightly more complex. This perceived complexity was associated with more iterative and in-depth engagement with both GenAI-generated critical reviews and the original academic article, accompanied by an increased tendency to validate and assess information across multiple sources due to concerns about the reliability of GenAI-generated content, a pattern also reported by Xiao and Mao. Deeper-level planning strategies appeared in their prioritization of purposes and research design, and in their pre-planning of how they would evaluate the GenAI-generated content against their expectations, according to Li. Deeper-level monitoring strategies were evident in learners' continuous efforts to check their understanding of the theoretical frameworks through comparison with their own reviews and the academic article, as demonstrated by Xiao. Deeper-level evaluation strategies reinforced this iterative process, with learners assessing the credibility and criticality of GenAI content while consulting the academic article for accuracy, as seen in Mao's comments.

I plan to first examine the purposes and research design of this experiment, which is more complicated, and see if the results match my expectations (Li).

At first, I firmly believed that this article had no conceptual framework. However, the GenAI-generated critical review claimed it used a theoretical framework. I then realized that my understanding of a theoretical framework might differ from the AI's. So, I went back to the original academic paper to compare and verify whether it was actually there (Xiao).

Since the course statistical methods is only offered next semester, I don't fully understand some aspects of the research design and the details. As a result, I sometimes doubt my own writing and repeatedly check whether I've missed anything. I also question whether the information provided by the GenAI-generated critical review is accurate and whether its critiques are valid. But to verify these points, I still have to go back to the academic article (Mao).

Third, in Pattern 3, the fifth critical reading aspect, evaluating the adequacy of evidence to support the main arguments, was perceived as the most cognitively demanding task, as it required abstract reasoning and synthesis rather than simple reference to the article. To manage this challenge, beyond carefully reading GenAI-generated critical reviews to rewrite their own critical reviews, learners adopted deeper metacognitive strategies by engaging in a novel approach: interacting with the GenAI chatbot, which appeared to shift AI's role from a mere content provider to a learning partner that may foster higher-order thinking through debate and discussion, as reflected in Zhang's quote. Deep-level planning strategies were demonstrated in their purpose-setting to seek chatbot assistance due to the task's abstract nature, as Liang noted. Learners also applied deep-level monitoring strategies to track comprehension progress while reviewing difficult content, switching frequently between the GenAI-generated critical review and chatbot, as seen in Zhang's case. Deep-level evaluation strategies were especially strong, as learners assessed the chatbot's responses and refined prompt techniques to get more meaningful insights, as Liao reported.

The aspect of the adequacy of the evidence to support the main arguments is quite abstract and hard to directly find in the academic paper. So, I considered directly chatting with the chatbot to see if it could provide a clear explanation (Liang).

This aspect is difficult for me to understand, and the GenAI-generated critical review alone doesn't make it clear enough for me. As a result, I opened both the GenAI-generated critical review and the chatbot more frequently, reading them side by side, section by section, to help me understand it better (Zhang).

I feel like my way of asking questions might not be effective, as the GenAI chatbot only gave a conclusion. Actually, I want to know how much evidence is needed for a high-quality article, so I revised my prompts (Liao).

5 Discussion

5.1 Interaction patterns with GenAI-powered reading tools

This study identified three distinct interactive reading patterns during GenAI-supported critical reading among postgraduate students, showing how they selectively engaged with and coordinated different information sources across four of the ten aspects of academic critical reading. Pattern 1 involved interactions between self-written and GenAI-generated critical reviews, with a focus on contributions (aspect 2 of critical reading). Students entered the task with an intended focus for extracting information from the academic article and read GenAI-generated critical reviews without revisiting the article. Pattern 2 encompassed interactions between the self-written critical review and GenAI-generated critical reviews, as well as between self-written reviews and the academic article, focusing on aspects of purposes and research design (aspect 1), and the effectiveness of linking theoretical frameworks with their main arguments (aspect 6). Notably, Nguyen, et al. [32] observed a similar pattern: participants compared AI-generated content with their own writing while searching for relevant articles to verify information. However, their participants read articles for the first time during the task, which required more time for comparison. In contrast, our study incorporated a forethought stage, enabling students to set their focus in advance during GenAI-assisted reading. This advance focus-setting was intended to foster deeper critical thinking and support iterative cross-checking and revision decisions [41]. These two interactive patterns identified in our study were not observed by Cheung, et al. [11], likely because the participants in their study focused on evaluating ChatGPT-generated texts rather than academic articles and did not involve writing critical reviews, which was a reading purpose in our study, as reading purposes significantly influence students' use of reading strategies, leading to shifts in reading patterns [15, 42].

Pattern 3 featured direct communication with the GenAI chatbot, alongside similar interactions between self-written and GenAI-generated critical reviews on the adequacy of the evidence (aspect 5). Previous researchers like Peng, et al. [26] found that participants engaged in a critical reading pattern involving academic articles and a chatbot, differing from the present study because students repeatedly sought information from the article to answer chatbot-generated questions. This may be because the chatbot in their study lacked the ability to provide feedback, which hindered participants from assessing the accuracy of their answers and refining their prompts, thereby limiting reflection on interactions with the chatbot. Similarly, Altay, et al. [29] noted that the absence of personalized feedback in AI chatbots negatively affected participants' ability to reflect, while personalized feedback has been proven crucial for fostering goal-setting, self-evaluation, and reflection [1, 43, 44]. In our study, personalized feedback appeared to encourage a shift from passive strategy use when receiving GenAI-generated critical reviews to more active and agentic engagement: they initiated dialogue with the chatbot and exercised agency in steering the interaction and judging the chatbot's output. Behaviorally, students' selective use of the chatbot for the aspect of the adequacy of evidence was notable, suggesting a preference for targeted, conversational help-seeking when they felt less confident or were confused about this aspect. Beyond cognitive support, according to previous research, direct communication with a chatbot may also provide emotional and motivational support [45] by offering a low-stakes space to ask questions, reducing anxiety about "not knowing enough" or "lacking prior knowledge," and reassuring students that revisions are on the right track, although these affective benefits were not identified in our study and should be examined in future research.

5.2 Factors shaping learners' interaction patterns with GenAI-powered reading tools

The results indicated that the level of cognitive demand embedded in critical reading tasks was closely associated with students' reading patterns, with learners adapting their metacognitive strategies that shifted from surface-level to deeper, more analytical approaches in response to task complexity. Specifically, when learners judged a task as easy, they tended to adopt surface-level strategies, such as those involving recalling, paraphrasing content, or simply posing problems, corresponding to Pattern 1. In contrast, tasks perceived as challenging or problematic elicited deeper strategies, such as involving the synthesis and connection of ideas or the use of novel approaches, which underpin Patterns 2 and 3. This gradation echoes the findings of Schraw and Moshman [46] and Yang and Bai [24] that learners optimize metacognitive strategies to adapt to different reading conditions. Importantly, although all three patterns involve core metacognitive processes, planning, monitoring, and evaluation, their depth differs in line with Yang and Bai's [24] classification of perceived difficulty of task (easy and difficult) and Stanton, et al.'s [12] surface versus deep distinction.

Pattern 1, related to the perceived easiest aspect of critical reading, identifying an article's contributions (aspect 2), reflected primarily surface-level metacognitive strategies. This explains learners' minimal re-engagement with the academic article and their use of surface strategies, such as planning content recall and paraphrasing key points from the GenAI-generated critical review. This finding aligned with Pan, et al. [21], who identified that surface-level strategies might prevent learners from engaging deeply with more cognitive resources. However, deeper strategies, as discussed by Ku and Ho [25], go beyond merely recognizing problems and focus more on exploring solutions. Consequently, the perceived ease of the content may have discouraged learners from using deep strategies to generate new ideas or adopt novel approaches to regulate their reading behavior, might help account for their relatively limited engagement with multiple sources. Pattern 2 involved two challenging aspects of greater perceived difficulty: the purposes and research design (aspect 1), and the effectiveness of linking theoretical frameworks with main arguments (aspect 6). In this pattern, learners frequently employed deep metacognitive strategies by evaluating, verifying, and integrating perspectives from different sources. Previous research [47] indicated that the substantial cognitive effort learners invested in monitoring differences in understanding and evaluating GenAI-generated output facilitated the development of metacognitive strategies. This further encouraged them to validate information from both the GenAI-generated critical review and the academic article. Pattern 3, associated with the most cognitively demanding aspect of critical reading, evaluating the adequacy of evidence (aspect 5), was the only pattern involving direct communication with the GenAI chatbot with the use of deep metacognitive strategies. This novel approach aimed to enhance the explicitness and interactivity of their critical reading experience, aligning with Pan, et al. [21], who found that learners adopted deeper strategies during chatbot interactions. Moreover, the high cognitive load involved in critical reviewing may have led some learners to rely more heavily on external assistance from the chatbot or GenAI-generated critical reviews. Their limited cognitive resources might have been insufficient to sustain metacognitive strategies applied to the academic article, which could have influenced the reading patterns observed. This overshadowing effect, where competing information sources vie for cognitive resources and weaken attention to less dominant ones, has been documented in a previous study [48].

6 Implications and limitations

This study serves as a pioneering work that extends distinctions between surface and deep metacognitive strategies by examining how task complexity and cognitive demands influence students' selective use of these strategies and, in turn, shape their interactive reading patterns in a GenAI-assisted academic critical reading environment, thereby enriching our understanding of metacognition in AI-supported contexts and offering implications for pedagogy, research, and technology development. First, for educators, they can script a task workflow that position GenAI-generated content alongside chatbots, original texts, and other materials and requires students to critically evaluate, compare, triangulate, and justify their use of AI-generated suggestions, which can trigger students' deeper use of metacognitive strategies and facilitate their reading depth. Meanwhile, educators should emphasize meta-AI literacy (e.g., effective prompting, verification, and responsible attribution) and remind students that AI-generated content is a complementary resource rather than a replacement for critical thinking [49, 50], helping learners maintain ownership of their ideas and adhere to academic standards. Second, researchers are encouraged to further explore students' reading processes and metacognitive strategies to identify effective techniques that enhance reading skills and uncover challenges. Finally, AI developers could enhance chatbots by integrating additional digital affordances that promote deeper metacognitive strategies. For instance, incorporating more explicit deep-level metacognitive prompts could encourage students to use these strategies more effectively, thereby fostering critical thinking and supporting learners' competency development with AI.

Several limitations were identified. First, the study focused on L2 postgraduate learners' critical reading of articles in English education, with a participant sample that was mostly female, which might limit its generalizability, and whether the findings would differ in disciplines such as STEM remains unexplored. Second, because this study mainly relied on analyzing video recordings of students' behavioral data, and while interviews provided some qualitative insights, it was difficult to examine students' cognitive processes or divide behavioral codes with greater granularity; future studies are therefore suggested to adopt methods such as eye-tracking, keystroke logging, and stimulated recall to capture students' in-depth cognitive processes of reading. Last, the absence of a non-GenAI control group and separate experimental groups for different GenAI functions limited our ability to examine in detail how specific GenAI tools shape learners' interaction patterns. For example, a control group without AI support might reveal more surface-level metacognitive strategies, whereas comparing a chatbot group with a GenAI-generated review group could show whether different GenAI functionalities influence learners' reading depth and affective states, which future research should address.

Acknowledgements
We would like to express our sincere gratitude to all the students who participated in the study and the teachers who supported us throughout this work.

Author contributions
H.L. was responsible for data collection and analysis, wrote the main manuscript text, and contributed to manuscript revisions. W.W. supervised the project, conceptualized the study, developed the methodology, drafted the research proposal, and revised the manuscript. B.Z. assisted in research design, reviewed various sections of the manuscript, and assisted with manuscript editing. Y.Z. assisted in data analysis, provided valuable feedback and guidance throughout the writing process, and assisted with manuscript editing.

Funding
This work was supported by Macao Polytechnic University Grants (Grant ID: RP/FCA-07/2025) and Guangdong Basic and Applied Basic Research Foundation (2025A1515011744).

Data availability
Data supporting the findings of this study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
Ethical approval for this study was obtained from the Faculty of Applied Sciences, Macao Polytechnic University (Approval No. HEA003-FCA-2025). The study was also conducted in accordance with the Declaration of Helsinki. Written informed consent was obtained from all participants prior to their participation in the study. All procedures involving human participants were conducted in accordance with relevant guidelines and regulations.

Consent for publication
All participants gave informed consent for their data to be published.

Competing interests
The authors declare no competing interests.

Received: 10 July 2025 / Accepted: 16 March 2026
Published online: 31 March 2026

References

1. Shafiee Rad H. Reinforcing L2 reading comprehension through artificial intelligence intervention: refining engagement to foster self-regulated learning. Smart Learn Environ. 2025;12(1):23. https://doi.org/10.1186/s40561-025-00377-2.
2. Wu H, Liu W. Exploring mechanisms of effective informal GenAI-supported second language speaking practice: a cognitive-motivational model of achievement emotions. Discov Comput. 2025;28(1): 119. https://doi.org/10.1007/s10791-025-09635-w.
3. Xin JJ. Investigating EFL teachers' use of generative AI to develop reading materials: A practice and perception study. Language Teaching Research. 2024:13621688241303321. https://doi.org/10.1177/13621688241303321
4. Yıldız M, Kağan KH, Saadin O, et al. Can artificial intelligence identify reading fluency and level? Comparison of human and machine performance. Read Write Q. 2025;41(1):66–83. https://doi.org/10.1080/10573569.2024.2345593.

5. Barrot JS. Balancing innovation and integrity: an emerging technology report on SciSpace in academic writing. Technol Knowl Learn. 2025;30(1):587–92. https://doi.org/10.1007/s10758-024-09802-w.
6. Liu W, Wang Y. The effects of using AI tools on critical thinking in English literature classes among EFL learners: an intervention study. Eur J Educ. 2024;59(4):e12804. https://doi.org/10.1111/ejed.12804.
7. Van LH, Li CS, Wan R. Critical reading in higher education: a systematic review. Think Skills Creat. 2022;44:101028. https://doi.org/10.1016/j.tsc.2022.101028.
8. Din M. Evaluating university students' critical thinking ability as reflected in their critical reading skill: a study at bachelor level in Pakistan. Think Skills Creat. 2020;35:100627. https://doi.org/10.1016/j.tsc.2020.100627.
9. Lee Y-H, Jhang J-Y, Hong H-Y. Becoming epistemically active in online reading: facilitating elementary school students' multimodal multiple document reading via sourcing organizers. Comput Educ. 2024;216:105048. https://doi.org/10.1016/j.compedu.2024.105048.
10. Cromley JG, Kunze AJ, Dane AP. Multi-text multi-modal reading processes and comprehension. Learn Instr. 2021;71:101413. https://doi.org/10.1016/j.learninstruc.2020.101413.
11. Cheung KKC, Pun JKH, Li W. Students' holistic reading of socio-scientific texts on climate change in a chatGPT scenario. Res Sci Educ. 2024. https://doi.org/10.1007/s11165-024-10177-2.
12. Stanton JD, Sebesta AJ, Dunlosky J. Fostering metacognition to support student learning and performance. CBE Life Sci Educ. 2021;20(2):fe3. https://doi.org/10.1187/cbe.20-12-0289.
13. Wilson K. Critical reading, critical thinking: delicate scaffolding in English for academic purposes (EAP). Think Skills Creat. 2016;22:256–65. https://doi.org/10.1016/j.tsc.2016.10.002.
14. Le HV, Nguyen TAD, Le DHN, et al. Unveiling critical reading strategies and challenges: a mixed-methods study among English major students in a Vietnamese higher education institution. Cogent Educ. 2024;11(1):2326732. https://doi.org/10.1080/2331186X.2024.2326732.
15. He T-h. Reading for different goals: the interplay of EFL college students' multiple goals, reading strategy use and reading comprehension. J Res Read. 2008;31(2):224–42. https://doi.org/10.1111/j.1467-9817.2007.00355.x.
16. Zhang L, Zhang LJ. Relationships between Chinese college test takers' strategy use and EFL reading test performance: a structural equation modeling approach. RELC J. 2013;44(1):35–57. https://doi.org/10.1177/0033688212463272.
17. Bernardo ABI, Mante-Estacio MJ. Metacognitive reading strategies and its relationship with Filipino high school students' reading proficiency: insights from the PISA 2018 data. Humanit Soc Sci Commun. 2023;10(1):400. https://doi.org/10.1057/s41599-023-01886-6.
18. Zou X-L, Ou L. EFL reading test on mobile versus on paper: a study from metacognitive strategy use to test-media impacts. Educ Assess Eval Account. 2020;32(3):373–94. https://doi.org/10.1007/s11092-020-09320-0.
19. Phakiti A. A closer look at the relationship of cognitive and metacognitive strategy use to EFL reading achievement test performance. Lang Test. 2003;20(1):26–56. https://doi.org/10.1191/0265532203lt243oa.
20. Yao Y, Sun Y, Zhu S, et al. A qualitative inquiry into metacognitive strategies of postgraduate students in employing ChatGPT for English academic writing. Eur J Educ. 2024;n/a(n/a):e12824. https://doi.org/10.1111/ejed.12824.
21. Pan M, Lai C, Guo K. Effects of GenAI-empowered interactive support on university EFL students' self-regulated strategy use and engagement in reading. Internet High Educ. 2025;65:100991. https://doi.org/10.1016/j.iheduc.2024.100991.
22. Kwarikunda D, Schiefele U, Muwonge CM, et al. Profiles of learners based on their cognitive and metacognitive learning strategy use: occurrence and relations with gender, intrinsic motivation, and perceived autonomy support. Humanit Soc Sci Commun. 2022;9(1):337. https://doi.org/10.1057/s41599-022-01322-1.
23. Huang T-C, Chou Y-Y, Gao Z-Q, et al. Exploring the effects of the 6E instructional model on learners' metacognitive behaviors in 3D modeling. Int J Technol Des Educ. 2025. https://doi.org/10.1007/s10798-025-09972-3.
24. Yang C, Bai L. The use of metacognitive strategies by Chinese PhD students of social sciences in Australian universities. Int J Educ Res. 2019;97:43–52. https://doi.org/10.1016/j.ijer.2019.06.007.
25. Ku KYL, Ho IT. Metacognitive strategies that enhance critical thinking. Metacogn Learn. 2010;5(3):251–67. https://doi.org/10.1007/s11409-010-9060-6.
26. Peng Z, Liu Y, Zhou H, et al. CReBot: exploring interactive question prompts for critical paper reading. Int J Hum Comput Stud. 2022;167:102898. https://doi.org/10.1016/j.ijhcs.2022.102898.
27. 27. Yang ACM, Chen IYL, Flanagan B, et al. Automatic generation of cloze items for repeated testing to improve reading comprehension. Educ Technol Soc. 2021;24(3):147–158. https://doi.org/10.30191/ETS.202107_24(3).0011
28. Leonet O, Arocena E, Saragueta E. Young bilingual students' use of metacognitive strategies to overcome comprehension difficulties when reading in the foreign language. Literacy. 2025;59(2):135–50. https://doi.org/10.1111/lit.12383.
29. Altay S, Schwartz M, Hacquin A-S, et al. Scaling up interactive argumentation by providing counterarguments with a chatbot. Nat Hum Behav. 2022;6(4):579–92. https://doi.org/10.1038/s41562-021-01271-w.
30. Liu M, Zhang J, Nyagoga LM, et al. Student-AI question cocreation for enhancing reading comprehension. IEEE Trans Learn Technol. 2024;17:815–26. https://doi.org/10.1109/TLT.2023.3333439.
31. Toyokawa Y, Horikoshi I, Majumdar R, et al. Challenges and opportunities of AI in inclusive education: a case study of data-enhanced active reading in Japan. Smart Learn Environ. 2023;10(1):67. https://doi.org/10.1186/s40561-023-00286-2.
32. Nguyen A, Hong Y, Dang B, et al. Human-AI collaboration patterns in AI-assisted academic writing. Stud High Educ. 2024;49(5):847–64. https://doi.org/10.1080/03075079.2024.2323593.
33. Pan M, Guo K, Lai C. Using artificial intelligence chatbots to support English-as-a-foreign language students' self-regulated reading. RELC J. 2024;00336882241264030. https://doi.org/10.1177/00336882241264030.
34. Wallace M, Wray A. Critical Reading and Writing for Postgraduates. SAGE; 2021.
35. DiBenedetto MK, Zimmerman BJ. Construct and predictive validity of microanalytic measures of students' self-regulation of science learning. Learn Individ Differ. 2013;26:30–41. https://doi.org/10.1016/j.lindif.2013.04.004.
36. Zhang H, Yu L, Cui Y, et al. Mining classroom observation data for understanding teacher's teaching modes. Interact Learn Environ. 2022;30(8):1498–514. https://doi.org/10.1080/10494820.2020.1734029.
37. van Braak M, van de Pol J, Poorthuis AMG, et al. A micro-perspective on students' behavioral engagement in the context of teachers' instructional support during seatwork: sources of variability and the role of teacher adaptive support. Contemp Educ Psychol. 2021;64:101928. https://doi.org/10.1016/j.cedpsych.2020.101928.

38. Chen C-H, Chang C-L. Effectiveness of AI-assisted game-based learning on science learning outcomes, intrinsic motivation, cognitive load, and learning behavior. Educ Inf Technol. 2024;29(14):18621–42. https://doi.org/10.1007/s10639-024-12553-x.
39. Shang J, Huang Y, Xu M, et al. Competition between human learners and ChatGPT: enhancing university EFL students' reading comprehension and critical thinking through competitive questioning. Comput Assist Lang Learn. 2025;1–35. https://doi.org/10.1080/09588221.2025.2577356.
40. de Kleijn R, Van Leeuwen A. Reflections and review on the audit procedure: guidelines for more transparency. Int J Qual Methods. 2018;17(1):1609406918763214. https://doi.org/10.1177/1609406918763214.
41. Untoro J, Childs R, Bose I, et al. Tools to improve planning, implementation, monitoring, and evaluation of complementary feeding programmes. Matern Child Nutr. 2017;13(S2):e12438. https://doi.org/10.1111/mcn.12438.
42. Shih Y-C, Reynolds BL. The effects of integrating goal setting and reading strategy instruction on English reading proficiency and learning motivation: a quasi-experimental study. Appl Linguist Rev. 2018;9(1):35–62. https://doi.org/10.1515/applirev-2016-1022.
43. Krenn B, Würth S, Hergovich A. The impact of feedback on goal setting and task performance. Swiss J Psychol. 2013;72(2):79–89. https://doi.org/10.1024/1421-0185/a000101.
44. Wei W, Cheong CM, Zhu X, Lu Q. Comparing self-reflection and peer feedback practices in an academic writing task: a student self-efficacy perspective. Teach High Educ. 2024;29(4):896–912. https://doi.org/10.1080/13562517.2022.2042242.
45. Liu C-C, Chiu CW, Chang C-H, et al. Analysis of a chatbot as a dialogic reading facilitator: its influence on learning interest and learner interactions. Educ Technol Res Dev. 2024;72(4):2103–31. https://doi.org/10.1007/s11423-024-10370-0.
46. Schraw G, Moshman D. Metacognitive theories. Educ Psychol Rev. 1995;7(4):351–71. https://doi.org/10.1007/BF02212307.
47. Teng MF. Metacognitive awareness and EFL learners' perceptions and experiences in utilising ChatGPT for writing feedback. Eur J Educ. 2025;60(1):e12811. https://doi.org/10.1111/ejed.12811.
48. Yang J, Li Y, Calic G, et al. How multimedia shape crowdfunding outcomes: the overshadowing effect of images and videos on text in campaign information. J Bus Res. 2020;117:6–18. https://doi.org/10.1016/j.jbusres.2020.05.008.
49. Lin H, Wei W, Liu F. GenAI-assisted critical reading report revision: a mixed-methods study. ECNU Rev Educ. 2026;9(1). https://doi.org/10.1177/20965311251387350.
50. Lin H, Chen Z, Wei W, Lu H. GenAI tools in academic reading: a study on AI-assisted metacognitive strategies and emotional reactions. In: Cheng ECK, editors. Innovating Education with AI. AETS 2024. Lecture Notes in Educational Technology. Springer, Singapore; 2025. https://doi.org/10.1007/978-981-96-4952-5_7.

Publisher's note