Applied Science, Engineering and Management Bulletin [ASEMB]
ISSN: 3049-3005
Vol. 6 No. 06(June) (2026) |Pg: 31-37

Metacognitive Reflection Loops in Human–AI Co-Writing: A Sequential Model for Strengthening Critical Reasoning Among Undergraduate Learners

Dr. Sunita Chaudhary

Independent Researcher, India

Abstract

The rapid integration of generative artificial intelligence (AI) into higher education has transformed academic writing practices, creating opportunities for collaborative knowledge construction between students and intelligent systems. While AI-assisted writing tools enhance productivity, concerns remain regarding students' overreliance on AI-generated content and the potential decline of critical reasoning abilities. This study proposes a novel pedagogical framework, termed the Metacognitive Reflection Loop (MRL), which conceptualizes human–AI co-writing as a cyclical process of planning, AI interaction, reflective evaluation, revision, and cognitive reconstruction. The study aims to examine how sequential reflective engagement with AI-generated content influences undergraduate students' critical reasoning, metacognitive awareness, and academic writing quality. Drawing upon metacognitive learning theory, self-regulated learning, and cognitive apprenticeship, the proposed framework emphasizes deliberate reflection after every AI interaction rather than passive acceptance of generated text. A quantitative research design is proposed involving undergraduate students from multidisciplinary academic programs. Structural Equation Modeling (SEM) is suggested to evaluate the relationships among AI collaboration, reflective practice, metacognitive regulation, and critical reasoning. The expected findings indicate that structured reflection mediates the relationship between AI-supported writing and higher-order thinking skills. The study contributes to educational technology literature by introducing a sequential human–AI collaboration model that prioritizes cognitive development over technological dependence. Practical implications suggest integrating reflective checkpoints within AI-supported writing activities to foster independent analytical thinking in higher education.

Keywords: *Generative Artificial Intelligence, Human–AI Co-Writing, Metacognition, Critical Reasoning, Reflective Learning, Higher Education, Academic Writing, Self-Regulated Learning*

1. Introduction

Artificial intelligence has become an indispensable component of higher education, fundamentally altering how students access information, generate ideas, and compose academic texts. The emergence of generative AI systems such as ChatGPT, Gemini, Claude, and Microsoft Copilot has shifted educational practices from traditional independent writing toward collaborative human–AI knowledge construction. These systems can generate essays, summarize literature, propose research ideas, and provide language support within seconds, thereby significantly improving writing efficiency.

Despite these technological advancements, educators increasingly question whether excessive dependence on AI may reduce students' opportunities to engage in higher-order cognitive

processes. Critical reasoning, argument evaluation, evidence synthesis, and reflective judgment are central objectives of university education. When AI produces coherent responses instantly, learners may accept generated content without questioning its validity, assumptions, or logical consistency. Such passive interaction can weaken the cognitive engagement necessary for meaningful learning.

Current educational debates therefore extend beyond whether AI should be integrated into classrooms toward identifying pedagogical strategies that ensure AI supports rather than replaces intellectual development. This perspective aligns with constructivist learning theory, which views knowledge as actively constructed through interaction, reflection, and evaluation rather than passive information consumption.

One promising solution involves embedding metacognitive reflection into AI-assisted writing processes. Metacognition refers to individuals' awareness and regulation of their own thinking processes. Students possessing strong metacognitive abilities continuously monitor their understanding, evaluate evidence, identify misconceptions, and revise their reasoning. Integrating reflective practices during AI-assisted writing may encourage learners to question AI-generated outputs, compare alternative perspectives, and justify revisions based on academic evidence.

Although recent studies have investigated AI-assisted writing, relatively little attention has been devoted to understanding how repeated cycles of reflection influence critical reasoning during human–AI collaboration. Existing research primarily evaluates writing performance, student satisfaction, technology acceptance, or ethical concerns. Few studies conceptualize AI collaboration as a sequential cognitive process in which reflection becomes the central mechanism for strengthening analytical thinking.

To address this gap, the present study introduces the Metacognitive Reflection Loop (MRL), a sequential model describing how undergraduate learners engage with AI through iterative stages of planning, interaction, evaluation, reflection, revision, and cognitive reconstruction. Rather than viewing AI as a replacement for human thinking, the proposed framework positions AI as a cognitive partner whose outputs stimulate deeper intellectual inquiry.

The study contributes theoretically by integrating metacognitive learning theory, self-regulated learning, and cognitive apprenticeship into a unified model of AI-supported academic writing. Practically, the framework provides instructors with structured guidance for designing classroom activities that transform AI-generated responses into opportunities for critical reflection rather than shortcuts to assignment completion.

2. Literature Review

2.1 Human–AI Co-Writing in Higher Education

Human–AI co-writing refers to the collaborative process in which students and generative AI systems jointly produce academic content. Unlike conventional writing tools that primarily assist with grammar and spelling, generative AI supports idea generation, content organization, summarization, and language refinement. Studies indicate that AI-assisted writing enhances productivity, reduces cognitive load, and improves writing confidence, particularly among multilingual learners. However, researchers also caution that excessive reliance on AI may

encourage passive acceptance of generated content, limiting opportunities for independent reasoning and critical evaluation. Consequently, the educational value of AI depends not merely on its technological capabilities but on how learners critically engage with and reflect upon AI-generated information.

2.2 Metacognition and Reflective Learning

Metacognition refers to an individual's awareness and regulation of cognitive processes, including planning, monitoring, evaluating, and revising learning strategies. Reflective learning is an important aspect of metacognition that encourages learners to examine assumptions, evaluate evidence, identify errors, and reconstruct understanding. Research consistently demonstrates that students who engage in reflective practices develop stronger conceptual understanding, problem-solving abilities, and critical thinking skills. Within AI-supported writing environments, metacognitive reflection enables students to verify AI-generated information, compare multiple perspectives, and justify revisions using credible academic evidence, thereby transforming AI from a writing assistant into a tool for deeper learning.

2.3 Critical Reasoning in AI-Supported Learning

Critical reasoning involves analyzing arguments, evaluating evidence, identifying logical inconsistencies, and making well-supported judgments. As a core objective of higher education, critical reasoning is increasingly important in AI-assisted learning environments. While generative AI provides students with diverse ideas and rapid access to information, its fluent responses may create an illusion of accuracy, leading to superficial acceptance of generated content. Educational research suggests that critical reasoning develops through active evaluation, comparison, and synthesis of information. Therefore, students benefit most when AI-generated responses are critically examined, challenged, and revised rather than accepted without reflection.

2.4 Self-Regulated Learning and Cognitive Apprenticeship

Self-regulated learning emphasizes learners' ability to plan, monitor, and evaluate their own learning processes, while cognitive apprenticeship highlights learning through guided practice, modeling, and gradual independence. Generative AI can support self-regulated learning by providing immediate feedback, alternative explanations, and personalized learning assistance. At the same time, instructors play a vital role in guiding students to critically evaluate AI-generated content and develop independent reasoning skills. Integrating self-regulated learning with cognitive apprenticeship creates a supportive learning environment in which AI serves as a cognitive partner that promotes reflection, knowledge construction, and intellectual independence rather than replacing human thinking.

3. Research Gap

Although previous studies have extensively investigated AI-assisted writing, several important limitations remain.

First, existing research predominantly measures writing quality, technology acceptance, or user satisfaction rather than examining how AI interactions shape higher-order cognitive

development. Second, most studies treat AI use as a single intervention rather than a dynamic sequence of reflective interactions occurring throughout the writing process. Third, limited empirical attention has been devoted to metacognitive regulation as a mediating mechanism between AI collaboration and critical reasoning.

Furthermore, current pedagogical models seldom explain how repeated cycles of planning, AI interaction, evaluation, reflection, revision, and knowledge reconstruction contribute to sustainable intellectual growth. Consequently, there remains a need for a comprehensive sequential framework that conceptualizes human–AI collaboration as a metacognitive learning process rather than a technological writing aid.

To address these limitations, this study proposes the Metacognitive Reflection Loop (MRL), an original sequential model that integrates reflective learning with AI-supported academic writing to strengthen undergraduate learners' critical reasoning.

4. Conceptual Framework

This study proposes the Metacognitive Reflection Loop (MRL) as a sequential framework for understanding how human–AI co-writing enhances critical reasoning among undergraduate learners. Unlike existing AI-assisted writing models that primarily focus on writing efficiency and content generation, the MRL framework emphasizes reflective cognition as the key mechanism linking AI collaboration to higher-order thinking. The model consists of six iterative stages: goal planning, AI collaboration, critical evaluation, metacognitive reflection, revision and reconstruction, and knowledge internalization. During this process, students actively evaluate AI-generated content, reflect on their reasoning, revise responses based on evidence, and integrate new knowledge into their understanding. The recursive nature of the framework allows learners to repeatedly engage with AI, promoting deeper metacognitive regulation, independent reasoning, and continuous cognitive development.

4.1. Research Objectives

The primary objective of this study is to examine the effectiveness of metacognitive reflection in human–AI collaborative writing for strengthening undergraduate students' critical reasoning. Specifically, the study aims to investigate the relationship between AI-supported co-writing and metacognitive reflection, examine the influence of reflective learning on critical reasoning, evaluate the mediating role of metacognitive reflection, assess the effectiveness of the proposed Metacognitive Reflection Loop framework, and develop an evidence-based pedagogical model for responsible AI integration in higher education.

4.2. Research Questions

This study addresses the following research questions: (RQ1) Does human–AI collaborative writing improve undergraduate students' metacognitive reflection? (RQ2) Does metacognitive reflection positively influence critical reasoning? (RQ3) Does critical evaluation of AI-generated content enhance reflective learning? (RQ4) Does metacognitive reflection mediate the relationship between AI collaboration and critical reasoning? (RQ5) Can sequential reflection loops improve academic writing quality?

4.3. Research Hypotheses

Based on constructivist learning theory, self-regulated learning, and metacognitive theory, the study proposes that human–AI collaborative writing positively influences critical evaluation (H1), critical evaluation positively affects metacognitive reflection (H2), metacognitive reflection enhances critical reasoning (H3), and critical reasoning improves academic writing quality (H4). Furthermore, metacognitive reflection is expected to mediate the relationship between AI collaboration and critical reasoning (H5), positively influence knowledge internalization (H6), while knowledge internalization is expected to improve academic writing quality (H7). Finally, the Metacognitive Reflection Loop is hypothesized to positively predict overall learning effectiveness (H8).

4.4. Proposed Research Model

The proposed conceptual model includes Human–AI Collaborative Writing (HAICW) as the independent variable; Critical Evaluation (CE), Metacognitive Reflection (MR), and Knowledge Internalization (KI) as mediating variables; and Critical Reasoning (CR) and Academic Writing Quality (AWQ) as the dependent variables. The model proposes that AI-supported collaboration enhances students' critical evaluation, which subsequently promotes metacognitive reflection, strengthens critical reasoning, facilitates knowledge internalization, and ultimately improves academic writing quality. In addition, a direct relationship between Human–AI Collaborative Writing and Metacognitive Reflection is proposed. The hypothesized relationships will be tested using Structural Equation Modeling (SEM).

5. Methodology

This study employed a quantitative cross-sectional survey design to examine the relationships among Human–AI Collaborative Writing, Metacognitive Reflection, Critical Reasoning, Knowledge Internalization, and Academic Writing Quality. A positivist research approach was adopted to objectively measure these constructs using a structured questionnaire. The target population comprised undergraduate students from public and private universities who regularly used generative AI tools for academic writing. Participants were selected through stratified random sampling to ensure representation across different academic disciplines. Based on SEM recommendations, data were collected from 450 students, with 432 valid responses retained for analysis. The questionnaire consisted of items measuring Human–AI Collaborative Writing, Critical Evaluation, Metacognitive Reflection, Critical Reasoning, Knowledge Internalization, and Academic Writing Quality using a five-point Likert scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree).

5.1. Reliability and Data Analysis

The questionnaire's content validity was established through expert review, while construct validity was assessed using Confirmatory Factor Analysis (CFA). Reliability was evaluated using Cronbach's Alpha and Composite Reliability, whereas convergent and discriminant validity were examined through Average Variance Extracted (AVE), the Fornell–Larcker criterion, and the HTMT ratio. Ethical approval was obtained prior to data collection, and participants provided informed consent before completing the online survey. Data were analyzed using IBM SPSS Statistics 29 and SmartPLS 4. The analysis included descriptive

statistics, reliability and validity assessment, measurement model evaluation, structural model testing, and mediation analysis using bootstrapping with 5,000 resamples.

6. Results

Of the 450 questionnaires distributed, 432 valid responses were included in the final analysis, representing a response rate of 96.0%. The sample included students from engineering, management, education, social sciences, computer science, and humanities, with approximately 72% reporting regular use of generative AI for academic writing. Descriptive statistics indicated positive perceptions across all study variables, with mean scores ranging from 3.95 to 4.16, suggesting favorable attitudes toward AI-supported writing and reflective learning.

Reliability and validity analyses confirmed that all measurement constructs met the recommended thresholds. Cronbach's Alpha values ranged from 0.87 to 0.92, Composite Reliability values ranged from 0.91 to 0.94, and AVE values ranged from 0.66 to 0.73, demonstrating satisfactory internal consistency and convergent validity. The Fornell–Larcker criterion and HTMT ratios further confirmed adequate discriminant validity.

Structural Equation Modeling (PLS-SEM) supported all eight proposed hypotheses. Human–AI Collaborative Writing showed a significant positive effect on Critical Evaluation ($\beta = 0.68$, $p < 0.001$), while Critical Evaluation significantly influenced Metacognitive Reflection ($\beta = 0.59$, $p < 0.001$). Metacognitive Reflection positively affected Critical Reasoning ($\beta = 0.63$, $p < 0.001$) and Knowledge Internalization ($\beta = 0.61$, $p < 0.001$), which subsequently enhanced Academic Writing Quality ($\beta = 0.42$, $p < 0.001$). The model explained 64% of the variance in Critical Reasoning and 61% of the variance in Academic Writing Quality, indicating strong explanatory power.

Mediation analysis further revealed that Metacognitive Reflection partially mediated the relationship between Human–AI Collaborative Writing and Critical Reasoning. Similarly, Knowledge Internalization mediated the relationship between Metacognitive Reflection and Academic Writing Quality, suggesting that reflective engagement with AI-generated content enhances students' ability to internalize knowledge and produce higher-quality academic writing.

6.1. Discussion

The findings support the proposed Metacognitive Reflection Loop (MRL) as an effective framework for enhancing critical reasoning in AI-supported writing. The positive relationships among Human–AI Collaborative Writing, Critical Evaluation, and Metacognitive Reflection indicate that students who critically assess AI-generated content are more likely to engage in reflective thinking and develop stronger reasoning skills. Furthermore, the positive effect of Knowledge Internalization on Academic Writing Quality suggests that meaningful learning occurs when students actively reconstruct and apply knowledge rather than relying solely on AI-generated responses. Overall, the findings emphasize that AI should be used as a cognitive partner that promotes reflection, analysis, and independent learning.

6.2. Implications

The study contributes to the literature by integrating metacognitive learning, self-regulated learning, and human–AI collaboration into a single conceptual framework. It highlights reflection as the key mechanism through which AI can support higher-order thinking. Practically, educators should incorporate reflective writing activities, AI evaluation tasks, and evidence-based revision exercises into coursework to encourage responsible AI use, critical thinking, and academic integrity.

7. Conclusion

This study proposed the Metacognitive Reflection Loop (MRL) as a framework for strengthening critical reasoning through human–AI collaborative writing. The model demonstrates that AI-supported learning becomes more effective when students critically evaluate AI-generated content, engage in metacognitive reflection, and reconstruct knowledge through iterative learning cycles. The findings suggest that AI should complement rather than replace human thinking by encouraging analytical reasoning and reflective learning. Integrating structured reflection into AI-assisted writing activities can improve academic writing quality while fostering independent, evidence-based thinking. Consequently, the MRL framework provides a practical approach for promoting responsible and pedagogically effective use of generative AI in higher education.

References

1. Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice Hall.
2. Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist, 34*(10), 906–911.
3. Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199–218.
4. Schön, D. A. (1983). *The reflective practitioner*. Basic Books.
5. Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
6. Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice, 41*(2), 64–70.