

**Human-AI Collaborative Learning Media in English Language Education:
Effects on Writing Quality, Creativity, and Cognitive Engagement**

Haris Dibdyaningsih¹ and Hendra Sudarso²

¹STKIP Al Hikmah Surabaya, ²Akademi Penerbang Banyuwangi

¹Surabaya, Indonesia and ²Banyuwangi, Indonesia

harisdibdyaningsih@gmail.com and hendrasudarso27@gmail.com

Abstract: *The integration of artificial intelligence into educational settings has fundamentally transformed the landscape of language acquisition, shifting the paradigm from self-learning to dynamic human-AI collaboration. However, the exact impact of collaborative learning media on specific pedagogical outcomes, such as writing quality, creative expression, and sustained cognitive engagement, remains a complex area of inquiry. This paper presents a comprehensive literature review and proposes a theoretical instructional framework to evaluate these collaborative dynamics within English language education. By synthesizing cross-disciplinary insights on human-AI teaming, cognitive mental models, and role adaptation, this work highlights the delicate balance between leveraging AI for productivity and ensuring profound pedagogical development.*

INTRODUCTION

The rapid advancement of artificial intelligence has permeated numerous domains, transforming how complex tasks are approached and executed. In the realm of English language education, AI-driven learning media have evolved from simple grammar checkers into sophisticated co-authors and conversational partners. Leveraging human expertise in tandem with these intelligent systems can be highly efficient and beneficial, yet the full extent of this teaming in educational contexts is still being uncovered (Islam et al., 2023). English language learners are increasingly interacting with generative models to brainstorm, draft, and revise their compositions. Consequently, it is imperative to understand how these human-AI collaborative environments influence not just the final textual output, but the underlying cognitive processes of the learners.

Despite the widespread adoption of AI tools, a clear problem definition is required to scope their educational efficacy. The primary challenge lies in determining whether AI collaboration genuinely enhances a student's fundamental writing quality and creative

thinking, or if it merely outsources cognitive effort to a machine. Education represents a highly specialized domain where standards of quality are meticulously scrutinized due to their lasting impact on a learner's developmental trajectory (Ren et al., 2024). Therefore, the scope of this review focuses specifically on three core dimensions: the measurable quality of the written text, the degree of original creativity exhibited, and the depth of the student's cognitive engagement during the learning process.

Currently, existing approaches to integrating AI in educational workflows are insufficient for several reasons. First, many contemporary systems prioritize task execution speed and automation over profound learning outcomes, which can paradoxically lead to a degradation in human accuracy and skill retention (Ren et al., 2024). Second, most existing frameworks rely on static AI roles that fail to dynamically adapt to the evolving mental models and specific situational needs of the learner (Huang et al., 2025) (Holstein & Satzger, 2025). When tools do not foster critical reflection, students may passively accept AI-generated content rather than actively engaging in the hermeneutic process of writing (Glinka & Müller-Birn, 2023).

To address these critical gaps, this paper synthesizes current research and proposes a structured pedagogical model. Specifically, the contributions of this paper are as follows:

1. First, it establishes a multidimensional taxonomy of human-AI collaboration that maps cross-disciplinary AI teaming concepts to the specific requirements of English language education.
2. Second, it proposes a novel triadic instructional framework designed to systematically evaluate and enhance writing quality, creativity, and cognitive engagement in ESL (English as a Second Language) learners.

LITERATURE REVIEW

To understand the nuanced effects of AI on educational outcomes, it is necessary to examine the broader landscape of human-AI collaboration. The literature can be organized into three primary categories: task performance trade-offs, cognitive and mental model development, and adaptive role personalization. Each of these subtopics offers critical insights that inform the application of AI in English language teaching.

Human-AI Teaming and Task Performance Trade-offs

The first category focuses on the core idea of how human-AI collaboration impacts efficiency versus accuracy in complex administrative and cognitive tasks. Research indicates that while AI approaches excel at specific rule-based operations, introducing human-in-the-loop (HitL) dynamics can yield performance that outpaces either human or AI agents acting autonomously (Islam et al., 2023). However, in educational contexts, this collaboration presents distinct strengths and weaknesses. A primary strength is the significant reduction in time required to complete tasks; for instance, human-AI collaboration can increase the speed of educational tagging tasks by approximately 50 percent (Ren et al., 2024). Conversely, a major weakness is that this increased speed often comes at the sacrifice of critical recall and accuracy, placing the combined performance below that of human-alone efforts in certain qualitative metrics (Ren et al., 2024). Compared to these general task-execution studies, the present work emphasizes that in language writing, the pedagogical goal is not speed, but the deliberate cultivation of language proficiency and creative thought.

Cognitive Engagement and Mental Models

The second category investigates the cognitive dimensions of human-AI interaction, emphasizing how human mental models evolve through continuous system use. The core idea here is that effective collaboration requires the alignment of domain knowledge, information processing, and complementarity-awareness (Holstein & Satzger, 2025). A notable strength of focusing on cognitive engagement is the promotion of "critical-reflective" collaboration, which is essential for scholarly knowledge production and meaningful interpretative tasks (Glinka & Müller-Birn, 2023). The weakness of this approach is that designing software that successfully forces a user to reflect, rather than merely consume, remains highly complex and difficult to standardize (Glinka & Müller-Birn, 2023). While previous research has explored these mental models in organizational decision-making and the humanities, this paper adapts these critical-reflective principles directly to the iterative process of English essay drafting.

Adaptive Roles and Personality Personalization

The third category explores the dynamic assignment of roles and personality traits to AI agents to optimize teamwork. Research proposes triadic frameworks where AI can dynamically shift between roles such as Advisor, Co-Pilot, and Guardian depending on

the user's immediate needs (Huang et al., 2025). Furthermore, large-scale studies have demonstrated that pairing specific human personality traits (e.g., extraversion, neuroticism) with customized AI agent personalities (e.g., conscientious, agreeable) significantly influences both the quality of the output and the success of the collaboration (Ju & Aral, 2025). The strength of this personalization is its ability to maximize engagement and tailor the collaboration to individual temperaments (Ju & Aral, 2025). The weakness is the substantial computational and ethical overhead required to profile learners and dynamically adjust AI behavior. This work integrates the concept of triadic role-shifting and personality matching to propose a more responsive English learning medium.

METHODOLOGY

To practically assess the impact of human-AI collaborative media on writing quality, creativity, and cognitive engagement, we propose a structured hypothetical framework: the Triadic Reflective Writing Framework (TRWF). This approach translates abstract human-AI interaction models into a concrete, phased pedagogical pipeline. The framework relies on the premise that effective human-AI collaboration should be distributed across distinct phases of the creative process, mirroring two-phase visualization systems where human-led and AI-mediated tasks are clearly delineated (Ouyang et al., 2024).

The TRWF operates through a numbered pipeline consisting of three distinct instructional modules:

1. **AI-Assisted Ideation (Advisor Role):** The student interacts with an AI to brainstorm topics and generate structural outlines. The AI acts purely as an advisor, providing divergent semantic prompts without generating complete sentences, thereby stimulating initial creativity.
2. **Co-creative Drafting (Co-Pilot Role):** As the student writes in English, the AI acts as a silent teammate or co-pilot, offering real-time vocabulary suggestions and syntactic corrections (Outmani et al., 2026). The AI utilizes a personality-paired interface, matching its feedback tone (e.g., encouraging vs. analytical) to the student's learning profile to maintain high cognitive engagement (Ju & Aral, 2025).
3. **Critical-Reflective Review (Guardian/Reviewer Role):** In the final phase, the traditional collaboration pattern is flipped to an "AI-creator + human-reviewer" dynamic (Li et al., 2025). The AI generates an alternative version of the student's

paragraph, and the student must critically evaluate both versions, justifying their linguistic choices to finalize the draft.

The key design choices in this framework are rooted in the necessity to prevent passive machine reliance. By restricting the AI from generating bulk text during the ideation and drafting phases, the cognitive load remains intentionally high, preserving the humanistic touch demanded by the educational enterprise (Ren et al., 2024). The rationale for the final reflective review phase is to enforce the development of domain mental models and complementary awareness (Holstein & Satzger, 2025). This ensures that the learner is actively analyzing grammatical structures and rhetorical devices, thereby securing long-term language acquisition.

To validate the TRWF, we propose a hypothetical evaluation plan involving a dataset of 200 intermediate ESL students over a 12-week semester. The students will be divided into a control group (using standard word processors) and an experimental group (using the TRWF platform). Writing quality will be evaluated using standardized IELTS grading rubrics assessed by independent, blinded educators. Creativity will be measured using divergent thinking metrics applied to the narrative structure of the essays. Finally, cognitive engagement will be quantified through think-aloud protocols and detailed interaction log data, analyzing the frequency and circumstances under which students exercise discernment by rejecting AI recommendations (Ren et al., 2024).

FINDING AND DISCUSSION

The proposed framework carries significant practical implications for the deployment of generative AI in language education. To ensure successful integration, institutions must consider the complex adoption factors of human-AI collaboration, recognizing that theoretical benefits do not automatically translate to user acceptance (Hemmer et al., 2022). Teachers and students must view the AI not as an infallible oracle, but as a coequal partner in the decision-making and writing process (Hemmer et al., 2022). Training programs should be established to help students build accurate mental models of the AI's capabilities and limitations, fostering an environment of active, rather than passive, learning (Holstein & Satzger, 2025).

Despite its potential, this approach presents several critical limitations and failure modes. First, there is an ever-present risk of over-reliance; if students prioritize task completion

speed over learning, human-AI collaboration may degrade their independent grammatical accuracy and recall (Ren et al., 2024). Second, measuring creativity remains highly subjective, and AI-assisted outputs might visually mimic creativity while actually resulting from formulaic algorithmic patterns. Third, context and cultural mismatch can occur; AI models trained on generalized datasets may fail to appropriately interpret or preserve the unique cultural nuances and idiomatic expressions intrinsic to an ESL learner's authentic voice.

Ethical considerations and risks must also be meticulously managed in these deployments. A primary ethical risk is the homogenization of student writing. If an entire cohort utilizes the same AI assistant, their distinct individual voices may converge toward an algorithmic mean, stripping the educational enterprise of its diverse, humanistic qualities (Ren et al., 2024). A second major risk revolves around data privacy and profiling. Implementing personality-paired AI requires the continuous tracking and analysis of a student's behavioral and psychological traits (Ju & Aral, 2025). Storing and processing such sensitive data raises profound consent and security issues that must be navigated with stringent governance.

Looking forward, future work should expand on the empirical foundations of collaborative learning media. One crucial direction is the execution of longitudinal studies that track ESL learners across multiple academic years to determine if AI-assisted skill gains translate into unassisted writing proficiency. Additionally, future research should explore cross-cultural adaptations of AI personality pairing, investigating how different demographic groups respond to varying levels of AI conscientiousness or extraversion during complex language tasks (Ju & Aral, 2025).

CONCLUSION

The integration of human-AI collaborative media in English language education represents a profound shift in pedagogical methodology. This literature review highlights that while AI can vastly improve the speed and efficiency of text generation, standard applications often threaten the critical accuracy, creativity, and cognitive engagement necessary for genuine language acquisition. By drawing upon cross-disciplinary frameworks of triadic role shifting, mental model development, and critical reflection, educators can design systems that prioritize learning over mere task execution.

Ultimately, the goal of AI in language education should not be to replace the student's cognitive effort, but to elevate it. The proposed Triadic Reflective Writing Framework serves as a foundational step toward achieving this balance. As AI systems continue to evolve from simple tools into interactive co-authors, continuous empirical evaluation and strict ethical governance will be vital to ensure that human-AI collaboration enriches the educational landscape, fostering writers who are both technically proficient and profoundly creative.

REFERENCE

- Islam, Md Saiful, Das, Srijita, Gottipati, Sai Krishna, Duguay, William, Mars, Clodéric, Arabneydi, Jalal, Fagette, Antoine, Guzdial, Matthew, & Matthew-E-Taylor (2023). *Human-AI Collaboration in Real-World Complex Environment with Reinforcement Learning*. <https://arxiv.org/pdf/2312.15160v1>
- Ren, Cheng, Pardos, Zachary, & Li, Zhi (2024). *Human-AI Collaboration Increases Skill Tagging Speed but Degrades Accuracy*. <https://arxiv.org/pdf/2403.02259v1>
- Huang, Gaojian, Jin, Yantong, & Lo, Wei-Hsiang (2025). *Beyond Levels of Driving Automation: A Triadic Framework of Human-AI Collaboration in On-Road Mobility*. <https://arxiv.org/pdf/2504.19120v1>
- Holstein, Joshua, & Satzger, Gerhard (2025). *Development of Mental Models in Human-AI Collaboration: A Conceptual Framework*. <https://arxiv.org/pdf/2510.08104v1>
- Glinka, Katrin, & Müller-Birn, Claudia (2023). *Critical-Reflective Human-AI Collaboration: Exploring Computational Tools for Art Historical Image Retrieval*. <https://arxiv.org/pdf/2306.12843v1>
- Ju, Harang, & Aral, Sinan (2025). *Personality Pairing Improves Human-AI Collaboration*. <https://arxiv.org/pdf/2511.13979v3>
- Ouyang, Yang, Zhang, Chenyang, Wang, He, Ma, Tianle, Jiang, Chang, Yan, Yuheng, Yan, Zuoqin, Ma, Xiaojuan, Shi, Chuhan, & Li, Quan (2024). *A Two-Phase Visualization System for Continuous Human-AI Collaboration in Sequelae Analysis and Modeling*. <https://arxiv.org/pdf/2407.14769v1>

Outmani, Moustapha El, Shenoy, Manthan Venkataramana, Hatahet, Ahmad, Rausch, Andreas, Kniep, Tim Niklas, Raddatz, Thomas, & King, Benjamin (2026). *Human-AI Collaboration for Scaling Agile Regression Testing: An Agentic-AI Teammate from Manual to Automated Testing*. <https://arxiv.org/pdf/2603.08190v2>

Li, Haotian, Wang, Yun, & Qu, Huamin (2025). *Reflection on Data Storytelling Tools in the Generative AI Era from the Human-AI Collaboration Perspective*. <https://arxiv.org/pdf/2503.02631v2>

Hemmer, Patrick, Schemmer, Max, Riefle, Lara, Rosellen, Nico, Vössing, Michael, & Köhl, Niklas (2022). *Factors that influence the adoption of human-AI collaboration in clinical decision-making*. <https://arxiv.org/pdf/2204.09082v1>