Generative AI as a Reflective Scaffold in a UAV-based STEM Project: A Mixed-Methods Study on Students' Higher-Order Thinking and Cognitive Transformation

ABSTRACT: In today's knowledge society, where artificial intelligence (AI) technologies are increasingly integrated into educational contexts, designing intelligent learning systems with cognitive scaffolding functions has become a critical issue in learning technology research. Unmanned Aerial Vehicle (UAV) courses, which involve mechanical assembly, sensor integration, data processing, and flight control, present a highly integrated STEM learning task and an ideal context for examining the effectiveness of generative AI-assisted learning. This study employed a quasi-experimental design involving 64 first-year engineering students, who were randomly assigned to either an experimental group or a control group to participate in a six-week UAV-based project-oriented STEM course. The experimental group interacted with a GPT-based system featuring semantic prompts and dynamic feedback to support reflective questioning and strategic adjustment, while the control group engaged in paper-based reflection activities. Independent sample t-tests revealed that students in the experimental group significantly outperformed their counterparts in STEM literacy, levels of reflection, and higher-order thinking indicators. Furthermore, thematic analysis of students' reflection records and AI dialogues demonstrated that the GPT system effectively facilitated verbalized reflection and multi-level cognitive regulation. Students' engagement gradually shifted from operational procedures to conceptually driven strategic understanding. This study confirms the feasibility of AI-supported reflection and provides concrete theoretical and practical implications for the design and evaluation of future intelligent learning systems.

Keywords: Generative AI, Reflective thinking, Higher-order thinking skills, UAV-based engineering learning, STEM education

1. Introduction

Contemporary STEM education has evolved from a subject integration approach aimed merely at enhancing technological literacy to a core pedagogical model that fosters students' higher-order thinking and self-directed learning capabilities. Global educational trends increasingly emphasize interdisciplinary knowledge integration and problem-based learning, advocating for students' application of knowledge in authentic contexts to develop logical, strategic, and creative problem-solving skills (Fajrina et al., 2020; Lee et al., 2023). Within this framework, the higher-order cognitive skills of Analyze, Evaluate, and Create are widely recognized as key indicators for assessing the effectiveness of STEM curricula. Despite efforts across Asia to promote STEM education, its practical implementation often encounters cultural and systemic barriers. Predominantly lecture-based instruction and a focus on standardized answers constrain students' capacity for critical thinking and self-questioning. Even when hands-on activities are included, students tend to remain at the level of rote procedural repetition, failing to engage in deeper understanding or adaptive strategy refinement (Abedi et al., 2023; Rasul et al., 2023). For instance, Taiwan's recent shift toward competency-based, interdisciplinary curricula aims to foster inquiry-driven and integrative learning. However, in practice, many educators lack experience in scaffolding higher-order thinking, and students are often deprived of semantic feedback mechanisms. Consequently, classroom activities frequently result in superficial collaboration and outputs, falling short of fostering meaningful cognitive processes (Jiang et al., 2024).

Unmanned Aerial Vehicles (UAVs), which integrate sensing, control, data processing, and mathematical modeling, have become increasingly popular as tools in interdisciplinary STEM instruction. Students must complete tasks such as module design and parameter calibration within a short timeframe while responding to anomalies during flight, thereby demanding not only technical competence but also causal reasoning, systems integration, and strategic planning (Diez et al., 2021; Yeni et al., 2024). UAV courses are thus regarded as cognitive pressure zones, where students lacking task structure and real-time feedback support are susceptible to operational frustration and cognitive overload (Chauhan & Sevda, 2023; Xue et al., 2025). These courses also exemplify Engineering Project-Based Learning (EPBL), which emphasizes the development of knowledge, strategies, and reflective adjustment through engagement with authentic engineering tasks. Effective learning in EPBL contexts is not limited to the production of final outcomes; rather, it involves the ability to self-interrogate, clarify concepts, and revise strategies during the learning process (Castro & Oliveira, 2023; Tritico & Korach, 2024). When students can articulate clear semantic explanations of their decisions and performance, they demonstrate learning that moves beyond knowledge application toward cognitive restructuring (Fu et al., 2025). However, in current teaching practices, instructors often struggle to deliver individualized and in-depth feedback in real time. As a result, students lack semantic guidance and incentives for reflection, which hinders deeper cognitive growth and leaves them confined to surface-level operations and factual recall (Abuhassna et al., 2022; Clarke & Konak, 2025).

Generative AI, with its capabilities in semantic generation, contextual understanding, and responsive dialogue, offers promising possibilities for addressing the limitations of teachers' cognitive feedback capacity. Unlike conventional structured learning tools, Generative AI can conduct semantic analysis based on students' inputs and provide timely strategic suggestions, guiding learners to formulate hypotheses, diagnose errors, and engage in self-regulation. It functions not merely as a question-answering assistant but as a linguistically responsive learning partner with process-tracking capabilities—thereby enabling dynamic cognitive scaffolding (Tran et al., 2025; Zawacki-Richter et al., 2019). However, current research predominantly explores the application of Generative AI in language learning, programming, or personalized education, while empirical investigations into its role in supporting reflection and higher-order thinking in engineering-based learning contexts remain scarce (Dwivedi et al., 2021). In cognitively demanding and time-constrained learning environments such as UAV courses, teachers often struggle to provide timely, strategic feedback. Consequently, students may find it difficult to produce meaningful verbal reflections, limiting both cognitive transformation and the development of higher-order thinking (AlAli, 2024). Without appropriate semantic scaffolding, learners tend to focus on procedural adjustments without advancing to analysis or evaluation levels. Reflective behaviors thus become superficial, undermining the learning potential of EPBL (Baričević & Luić, 2023).

In response to these challenges, this study proposes the use of Generative AI as a cognitive scaffold throughout the engineering task process. A GPT-based interactive system was developed to provide semantic prompts and dynamic feedback, supporting students in engaging in reflective inquiry, concept validation, and strategic adjustment within a UAV-centered, project-based STEM curriculum. Grounded in semantic analysis and process tracking, the system serves as a digital mediating tool that facilitates verbalized reflection and cognitive regulation, thereby exploring its functions and mechanisms in promoting students' higher-order thinking and learning transformation. This study also addresses a gap in current literature regarding the empirical analysis of AI-supported reflective processes in engineering learning environments. Accordingly, the study is guided by the following research questions:

- RQ1: Does the integration of Generative AI into an interdisciplinary engineering project-based curriculum significantly enhance students' STEM literacy?
- RQ2: In knowledge-integration tasks such as UAV design, does the semantic guidance provided by Generative AI promote students' reflective and higher-order thinking?
- RQ3: How does Generative AI support the development of students' verbalized reflection and strategic regulation during the learning process?

For clarity, Generative AI in this study refers broadly to AI systems with natural language generation and interactive reasoning capabilities, applied across tasks involving language comprehension, dialogue generation, and knowledge construction. The specific implementation uses a GPT model as the tool for semantic prompting and cognitive scaffolding. While other models such as Claude and Gemini fall under the same category, they are not included in the current analysis due to the study's focus on a single-tool application. Overall, the design emphasizes the role of GPT as a digital mediator in supporting linguistic reflection and strategic regulation throughout the instructional process.

2. Literature review

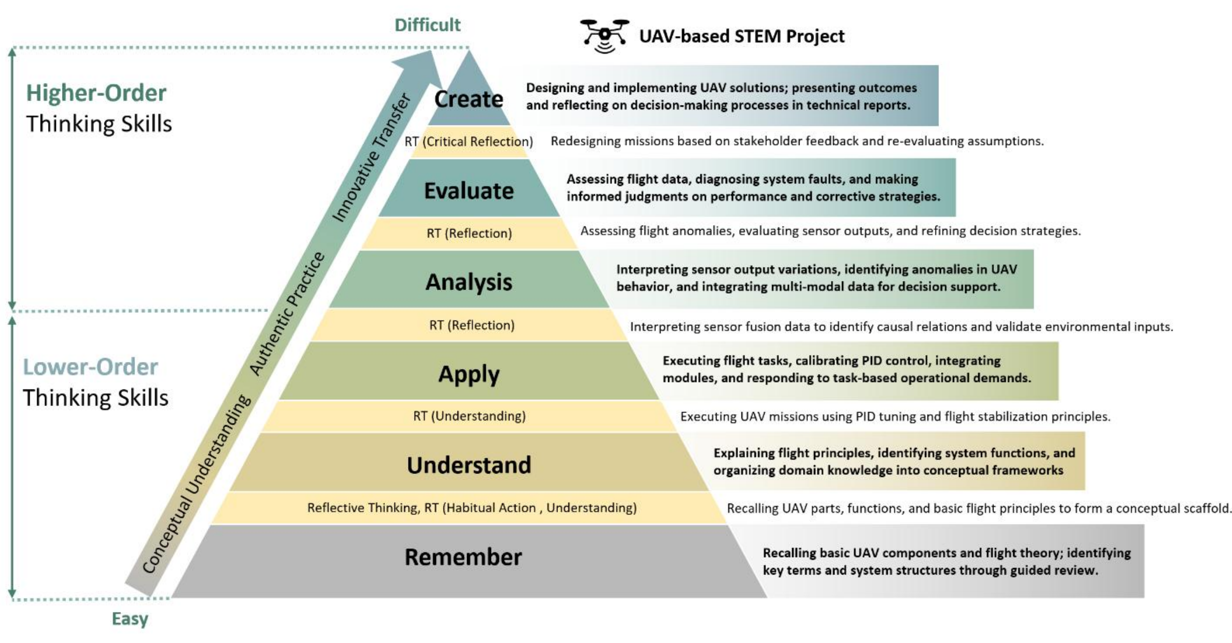
2.1. Theoretical foundations of higher-order cognition and reflective thinking in UAV-based STEM education

As modern education increasingly emphasizes the development of students' deep reflection and higher-order thinking, Bloom's Taxonomy remains a foundational framework for curriculum design and learning assessment (Bloom, 1964). This taxonomy originally classified cognitive processes into six hierarchical levels: Knowledge, Comprehension, Application, Analysis, Synthesis, and Evaluation. Anderson and Krathwohl (2001) later revised the taxonomy into a more action-oriented model. In this revised version, the levels of Analyze, Evaluate, and Create are identified as the core components of Higher-Order Thinking Skills (HOTS), which are frequently used in addressing open-ended problems and strategic learning tasks (Aldalur and Sagarna, 2023; Baričević and Luić, 2023). However, the development of HOTS is not limited to cognitive process operations. Dewey (1933) proposed the concept of reflective thinking, which emphasizes purposeful thought that arises through the evaluation and reconstruction of experiences. This process supports learners in building deeper understanding and fostering self-regulation. Kember et al. (2008) outlined four levels of reflection: habitual action, understanding, reflection, and critical reflection. These levels have been recognized as crucial elements in transformative learning (Clarke and Konak, 2025; Fu et al., 2025). Schön (1987) further introduced a dual framework consisting of reflection-in-action and reflection-on-action, which has served as a significant reference for the design of reflective curricula (Chauhan and Sevda, 2023).

STEM courses often present learners with complex, uncertain, and non-standardized problems that demand decision-making, strategic adjustment, and the integration of knowledge across multiple domains. These challenges place considerable demands on students' reflective thinking and higher-order cognition (Guo et al., 2025; Shehata et al., 2024). UAV-based project learning environments are typical examples of such tasks. These courses combine elements of flight control, sensor integration, data processing, and mathematical modeling, forming learning environments characterized by high time pressure and technical complexity (Shadiev and Yi, 2023; Yeni et al., 2024). Within these settings, students are required to document flight data, analyze errors, and revise their designs during practical tasks. These activities call for both the cognitive processes defined in HOTS and the metacognitive regulation involved in reflective thinking (Sukackė et al., 2022).

Recent research has shown the potential of UAV-based STEM curricula to support and deepen reflective learning. For example, Yeung et al. (2025) found that in a UAV water sampling project, students enhanced their learning motivation and self-monitoring abilities through strategic team collaboration and iterative task optimization. These processes also increased the depth of reflection. However, challenges remained in students' development of mathematical reasoning and logical explanation. Subramaniam et al. (2025) used the WoT4EDP framework to argue that engineering-oriented tasks integrating design thinking, scientific inquiry, and metacognitive reflection can help students develop more comprehensive thinking systems. To better understand how higher-order thinking and reflective processes develop in UAV-based STEM learning, this study integrates Bloom's revised taxonomy with levels of reflective thinking as an interpretive model. As illustrated in Figure 1, this model demonstrates how each cognitive level in Bloom's Taxonomy can be supported by increasing depths of reflection, ultimately fostering higher-order learning outcomes in STEM education.

Figure 1. Conceptual framework mapping Bloom's cognitive levels to corresponding depths of reflective thinking in a UAV-based STEM context



In addition, the integration of digital technologies has provided substantive support for students' reflective processes. Tools such as real-time feedback systems, video annotation platforms, and dynamic semantic prompting serve as effective digital scaffolds, enhancing students' sensitivity to reflection and adaptability in learning strategies (Abuhassna et al., 2022; Satar et al., 2024). These supports allow learners to co-develop higher-order cognition and reflective learning, particularly in the context of complex, open-ended tasks. For example, digital collaborative tools like Padlet, and reflective modules embedded in learning management systems (LMS), have been shown to facilitate structured verbal reflection and self-monitoring processes among students (AlAli, 2024). Wang et al. (2025) demonstrated that integrating Generative AI into virtual reality (VR) environments can enhance reflective thinking through personalized semantic feedback. This integration was also found to increase students' learning motivation and improve their AIoT implementation skills, highlighting the potential of generative AI to support deep learning in immersive settings. Similarly, Baričević and Luić (2023), through an augmented reality–based design thinking curriculum, found that actively engaging students in design processes significantly improved their innovative thinking and critical evaluation abilities. These findings suggest that the combination of immersive technologies with active learning strategies is effective in fostering higher-order thinking skills.

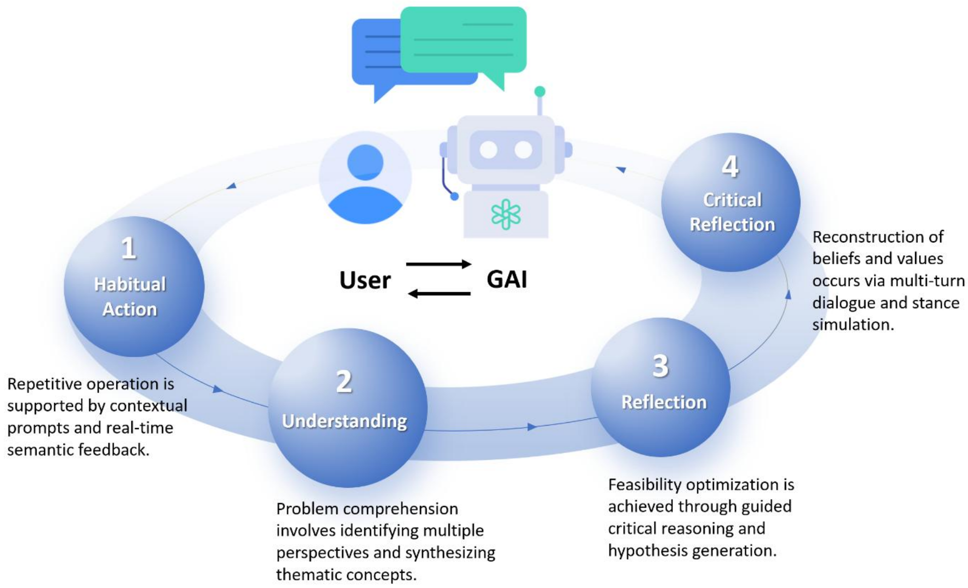
Overall, Bloom's Taxonomy and Reflective Thinking Theory form the theoretical foundation of higher-order learning, while UAV project-based education offers a concrete and practical implementation context. Through the integration of task-based learning and digital technology, students are guided to engage in decision-making, error correction, and conceptual restructuring. These processes collectively support the gradual development of reflective depth and higher-order thinking in STEM education.

2.2. Cognitive scaffolding roles of Generative AI in reflective thinking

Reflective thinking is not a momentary cognitive activity but a progressive and sustained learning process. Learners must continually revisit, evaluate, and reconstruct their experiences in order to refine prior understandings and deepen cognitive structures (Xue et al., 2025). Tutticci et al. (2018), in the context of high-fidelity simulation-based instruction, found that even when structured feedback was provided by instructors, students often struggled to achieve critical levels of reflection. This highlights the need for more responsive and flexible support mechanisms to foster deeper reflective engagement. As an emerging learning tool, Generative AI offers significant potential in addressing these challenges. It can help students identify cognitive blind spots, clarify conceptual misunderstandings, and provide structured semantic feedback that promotes depth in reflective thinking (Shi et al., 2023). Chen and Lin (2025) further observed that AI systems equipped with real-time questioning and reflection logging functions assist learners in becoming aware of their cognitive limitations, making strategic adjustments, and enhancing their learning self-regulation. Such capabilities are essential for advancing the reflective process by increasing learners' metacognitive awareness of knowledge gaps.

In practical applications, Generative AI can serve dual roles as both a semantic prompt provider and a generative dialogue partner. Semantic prompts help students articulate the reasoning behind their decisions, while generative dialogues introduce diverse perspectives, pose challenging questions, and simulate alternative scenarios. These features encourage conceptual restructuring and hypothesis testing (Kim et al., 2025; Yang et al., 2024). In language learning environments, Generative AI has been shown to co-construct semantic spaces with learners, acting as a dialogic partner that supports collaborative reflection and enhances learners' sensitivity to their own thinking processes (Zaim et al., 2025). The real-time feedback and exploratory semantic features of Generative AI align closely with Dewey's (1933) assertion that reflective thinking must be continuous and progressive. Yeung et al. (2024) emphasized that in open-ended tasks such as UAV design, AI-generated encouraging prompts and contextual cues can assist students in managing frustration and sustaining cognitive engagement. This illustrates the affective scaffolding role of Generative AI, which supports not only cognition but also the emotional regulation necessary for deep learning (Namaziandost et al., 2023). As shown in Figure 2, Generative AI interactions correspond with the four levels of reflective thinking proposed by Kember et al. (2008): habitual action, understanding, reflection, and critical reflection. These stages are supported through context-aware prompts, real-time feedback, guided reasoning, and dialogic simulation.

Figure 2. Generative AI-supported progression through reflective thinking phases



At the early stage of learning, students often remain in the phase of habitual action, where they tend to perform tasks without recognizing the underlying meanings of their behaviors. During this stage, generative AI can act as a cognitive activator, using semantic feedback and question prompts to guide learners toward the understanding phase. This helps students clarify their experiences and begin constructing meaning (Chen et al., 2023; Moon et al., 2024). For example, Fang et al. (2025) found in a debugging instructional context that AI support enhanced students' logical monitoring and problem analysis, thereby promoting strategic reflection. Once students reach the reflection stage, generative AI can offer tools such as summary synthesis, conceptual comparison, and strategy inference to help them evaluate the effectiveness of prior actions and make appropriate adjustments (Jauhari, 2024; Jiang et al., 2024; Gueye and Exposito, 2025). At this point, the value of AI lies not in information delivery but in facilitating critical reasoning and hypothesis generation. Lin et al. (2025) demonstrated that in design-based tasks, AI-supported guidance in storyboarding and self-questioning significantly aided novice learners in restructuring their perspectives and expanding their imaginative capacity. At the final stage of critical reflection, generative AI can simulate multi-turn dialogues involving contrasting positions, helping learners challenge their assumptions and values. This interaction facilitates deeper conceptual reorganization (Mezirow, 1991; Lievens, 2022). Meng et al. (2025) emphasized that, in teacher education programs, generative AI enabled preservice teachers to critically examine existing instructional beliefs and construct personalized teaching styles, thereby deepening instructional reflection. Similarly, Min et al. (2025) suggested that teachers embed AI as a virtual interlocutor within course design, enabling students' ideas to be constructively challenged from multiple perspectives and enhancing the process of transformative learning.

Despite the demonstrated potential of generative AI to support deep reflection and conceptual transformation, its powerful content generation capabilities have also raised significant concerns in educational settings. In particular, issues related to academic integrity, learning process traceability, and assessment design have made generative AI not merely a supplementary tool but a catalyst for rethinking instructional structures. Chaudhry et al. (2023), in a comparative study of ChatGPT and high-performing undergraduate students, highlighted the high efficiency of AI-generated content along with its implications for academic honesty. Their findings underscore the urgency for educational institutions to reconsider the design of assessments and cognitive engagement strategies. In summary, generative AI can function as a digital scaffold that provides semantic guidance, contextual prompting, and cognitive regulation throughout the reflective process. It supports learners in identifying confusion from experience, forming hypotheses, and reorganizing reasoning structures (Preiksaitis and Rose, 2023). This demonstrates its capability to facilitate multi-level reflection and conceptual transformation, positioning generative AI as a valuable reflective thinking support tool in the design of intelligent learning systems (Liu et al., 2023; Sandhaus et al., 2024; Yuan and Hu, 2024).

2.3. Educational applications of Generative AI in designing for strategic reflective learning

As Generative AI rapidly advances in educational contexts, its applications in curriculum design and strategic reflective instruction are becoming increasingly established. This is especially evident in areas such as enhanced learning interaction, concept visualization, and the generation of adaptive learning content (Lang et al., 2025). Educators and instructional designers are increasingly exploring how Generative AI can function as a scaffold for dynamic feedback and cognitive regulation to build personalized, structured, and scalable reflective environments. These environments aim to support the development of higher-order thinking and sustained behavioral optimization in learners (Çetinkaya et al., 2025; Chen et al., 2025).

Traditional approaches to reflective learning have relied on paper-based journals and face-to-face discussions. While these methods often provide depth, they lack immediacy and scalability (Ahmed et al., 2024). Generative AI, through its language generation capabilities, offers modular and context-aware feedback mechanisms that can help teachers develop reflective instructional strategies adaptable across various disciplines (Reddy et al., 2024). For example, in writing education, Generative AI can assist in identifying logical gaps, weak argumentation, and issues in paragraph cohesion (Jiao et al., 2024). Multi-perspective comparative analyses generated by AI can also enhance students' critical writing expression (Sarker et al., 2024). In programming education, Generative AI offers dialogic feedback to guide students through problem comprehension, strategy design, and debugging. Beginners using tools such as ChatGPT report reduced syntax errors, increased confidence, and greater learning motivation (Garcia, 2025). Its role as a collaborative learning partner has also been shown to improve students' computational thinking, self-efficacy, and engagement (Yilmaz et al., 2023).

Compared to language and computer science education, the use of Generative AI in STEM instruction is still emerging, particularly in high-precision, performance-dependent domains such as UAV learning. Nevertheless, growing research illustrates its potential. For example, Generative AI can simulate disaster response scenarios and provide semantic prompts without internet connectivity, supporting students in strategy evaluation and real-time decision-making (Kaleem et al., 2024). Hazarika et al. (2024) developed a multi-role simulation system in which learners discuss sensor deployment and airspace coordination with AI agents, thereby strengthening their abilities in task decomposition and problem formulation. Jiang et al. (2024) found that AI's contextual understanding and visual analysis capabilities can offer real-time suggestions for optimizing flight paths and reflecting on strategic decisions. Moreover, Generative AI demonstrates potential in UAV communication networks, supporting spectrum estimation, resource allocation, and information synthesis to meet the demands of complex mission environments (Sapkota et al., 2025; Sun et al., 2024). Zhou et al. (2025) proposed a terminal–edge collaborative AI framework that enables model decomposition and aggregation for UAV systems, balancing real-time inference with task flexibility and extending the applicability of AI in UAV instruction and operations.

In more authentic learning scenarios, Runnel et al. (2024) designed a multi-role dialogue simulation environment involving policymakers, residents, and scientists, allowing students to explore value conflicts and the consequences of action. This process promoted critical thinking and the ability to integrate multiple perspectives. Similarly, Fan et al. (2025), in a study on GAI-assisted pair programming, found that AI not only helped identify errors but also reduced students' anxiety, improved their strategic adaptability, and enhanced their sense of collaboration. Generative AI can also deconstruct complex tasks into multi-stage processes, supporting reflection cycles and action revisions at each stage (Holzinger et al., 2022). Overall, the integration of Generative AI expands the possibilities for reflective tools and challenges educators to rethink instructional strategies and the design of digital learning environments (Meneses, 2023). In high-risk, high-accuracy tasks such as UAV operations, AI-generated feedback can provide learners with clear and actionable guidance for cognitive adjustment, enhancing strategic evaluation and cognitive sensitivity. The pedagogical potential of Generative AI ultimately depends on educators' capacity to design and guide its use. This includes prompt design, data interpretation, and analysis of students' learning processes (Doğan & Şahin, 2024; Thararattanasuwan & Prachagool, 2024). Through modular instructional design and cross-disciplinary curriculum integration, educators can more effectively harness Generative AI's transformative potential in strategic reflective learning and help establish learner-centered intelligent learning ecosystems.

3. Method

3.1. Participants

Participants in this study were 64 first-year undergraduate students (N = 64) from the College of Engineering at a university in southern Taiwan. The experimental course was part of the institution's interdisciplinary STEM curriculum, focusing on UAV-based project implementation and systems integration. The course incorporated elements of problem-based learning and cross-domain engineering applications. A stratified random assignment procedure was used to divide students into an experimental group (EG, n = 32) and a control group (CG, n = 32), ensuring homogeneity in terms of prior academic background and baseline abilities. Students in the experimental group participated in a task-based learning environment supported by Generative AI. Through an interactive platform powered by a locally deployed GPT model, students engaged in reflective questioning, recorded learning processes, and completed guided reflective writing based on AI prompts. The control group received the same instructional content and hands-on UAV tasks but followed a traditional teaching model. Their reflection activities were completed using handwritten journals without the involvement of any AI technology. Prior to the course, all participants were fully informed about the study's purpose, procedures, and their rights as participants. Informed consent was obtained voluntarily from all students. The study strictly adhered to ethical principles of educational research, emphasizing participant autonomy, data confidentiality, and the right to withdraw at any time.

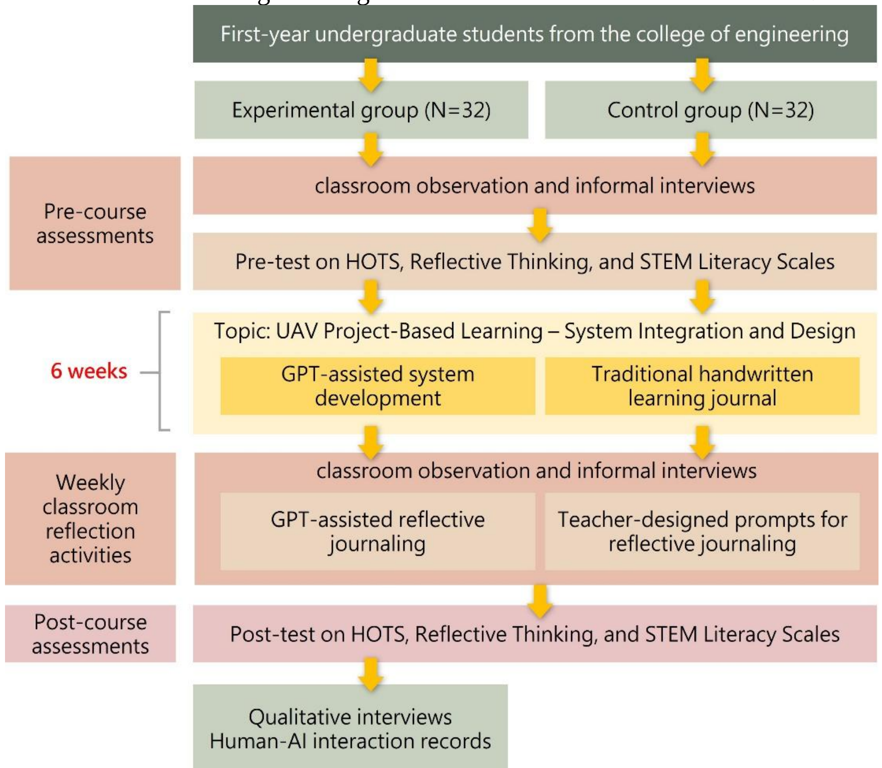
3.2. Experimental design

In this study, UAVs were not positioned solely as standalone technological tools but were conceptualized as integrative task carriers that embody the core competencies of STEM education. The course design incorporated elements from four disciplinary dimensions: science, technology, engineering, and mathematics. Scientific knowledge was embedded through the exploration of flight principles and sensor applications. Technological components involved the integration of hardware modules and the generation of operational data. Engineering tasks focused on mission planning and problem-solving strategies, while mathematical elements were reflected in data interpretation and the calibration of flight parameters. Through this interdisciplinary configuration, students engaged in a task-based learning process that required the application of multiple knowledge domains in authentic and cognitively demanding situations.

To support learners' cognitive and reflective engagement in such a complex learning environment, the course was enhanced with semantic prompting mechanisms powered by a Generative AI system. These mechanisms were designed to stimulate strategic thinking, guide reflective questioning, and promote conceptual transformation. By embedding AI-driven interactions into the course structure, the instructional design aimed to align with the pedagogical objectives of STEM education, particularly the development of real-world problem-solving abilities and higher-order cognitive processes such as analysis and evaluation.

The study adopted a quasi-experimental design, with participants randomly assigned into an experimental group and a control group. This design enabled the investigation of how the integration of Generative AI influenced students' learning performance in terms of Reflective Thinking and the HOTS. To strengthen the validity and depth of the findings, both quantitative and qualitative methods were employed. Pre-tests were administered to both groups prior to the course to assess baseline levels of STEM literacy, Reflective Thinking, and HOTS. In addition to these standardized instruments, classroom observations and informal interviews were conducted before instruction to collect contextual data about students' prior experience with UAVs and their initial perceptions of AI-supported learning. These insights were used to refine the design of system prompts and tailor the human-AI interaction model to the learners' backgrounds. The overall experimental procedure is illustrated in Figure 3, which outlines the sequential steps of instructional implementation, data collection, and analysis.

Figure 3. Research flowchart illustrating the integration of GPT-assisted reflection in a UAV-based STEM project



The instructional intervention lasted for six weeks, as outlined in Table 1. The curriculum was designed around UAV-based task learning, integrating multiple dimensions of STEM, including scientific principles, technological applications, engineering design, and mathematical modeling. All instructional activities were conducted in person, and the course was taught by a single instructor who possessed expertise in both UAV education and educational technology, thus controlling for instructional variability across groups. The course followed a two-phase structure. During the first phase, the instructor provided concept explanations, demonstrated system operations, and scaffolded core competencies through guided examples. In the second phase, students engaged in collaborative group work, autonomous exploration, and task execution. The experimental group utilized a custom-developed GPT-based interactive system to support reflective thinking. This system delivered multi-level prompts and dynamically adapted dialogue content based on each student's task progress. Students in the experimental group were required to interact with the GPT system at least three times per session. Two of these interactions were system-initiated prompts designed to guide reflection, while one was initiated by the student as an open-ended inquiry. The control group completed the same hands-on tasks and received identical content. However, their reflective activities were conducted on paper without any AI involvement. To ensure consistency, the instructor provided a standardized written prompt format focused on experience description and operational review. The control group did not receive personalized feedback or real-time language support to avoid confounding the validity of between-group comparisons.

Table 1. Weekly instructional plan for UAV-based STEM project learning

Week	Theme	S (Science)	T (Technology)	E (Engineering)	M (Mathematics)
Week 01	Introduction to UAV basics and flight principles	Understanding flight principles, gravity, and lift	Operating simulators; understanding module components (FC, motor, battery)	Exploring flight system architecture	Estimating basic aerodynamic forces
Week 02	Task scenario design and mission analysis	Investigating application domains (e.g., agriculture,	Retrieving sensor specifications and functions	Analyzing mission structure and defining goals and requirements	Estimating data requirements and computation scope

		environment)			
Week 03	Sensor selection and module integration	Testing environmental variables (temperature, pressure, distance)	Operating MCU and configuring data output	Designing sensor integration systems	Interpreting sensor data and assessing precision
Week 04	UAV assembly and flight control design	Calibrating flight stability and understanding parameters	PID tuning and system configuration	Assembling components and verifying flight simulator	Adjusting control parameters and analyzing effects
Week 05	Flight testing and fault diagnosis	Analyzing anomalies and sensor deviation	Diagnosing faults and processing data	Designing troubleshooting and error flowcharts	Analyzing flight trajectory and sensor logs
Week 06	Project presentation and learning reflection	Reviewing task execution and learning observations	Writing technical reports and reflecting on GPT dialogues	Presenting project outcomes and decision-making process	Synthesizing data use and final outcome evaluation

At the end of each weekly session, both groups were required to submit a reflective entry of 150 to 200 words. However, the structure and depth of these reflections were determined by the intervention type. Experimental group students were asked to expand on their reflections based on the dialogue content generated through GPT interactions, whereas control group students followed a fixed reflective template provided by the instructor. In both cases, students' reflections were expected to address the following three dimensions:

(1) Experience description (e.g., "Which feedback from the GPT interaction was most memorable for you this week?")
(2) Strategy evaluation (e.g., "How do you interpret the GPT system's response? Why did you choose this operational strategy?")
(3) Improvement projection (e.g., "How might the GPT system's suggestions influence your planning or flight design in future tasks?")

After the six-week course concluded, both groups completed post-tests using the HOTS Scale, the Reflective Thinking Scale, and the STEM Literacy Scale to assess changes in cognitive levels and reflective ability. In addition, semi-structured focus group interviews were conducted with students in the experimental group. These interviews were organized around three thematic dimensions:

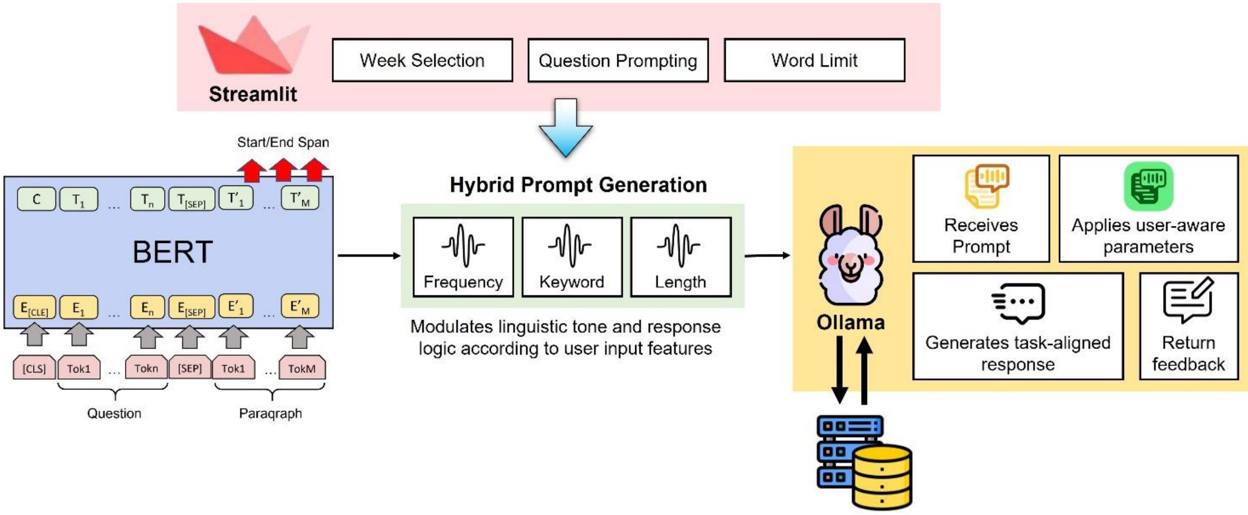
(1) perceptions of AI interaction and language-based feedback,
(2) awareness and development of reflective thinking levels, and
(3) adaptations in learning strategies and task planning.

To enhance the validity and consistency of the qualitative data, the interview sessions were supplemented by a review of GPT conversation logs and the students' weekly reflective entries.

3.3. System architecture

To support students' development of reflective thinking and HOTS during UAV-related project-based learning, this study developed a GPT-based interactive system that integrates reflective scaffolding strategies with large language model (LLM) technologies. The system comprises four major components: a front-end user interface, a semantic generation engine, a task-prompt logic module, and a corpus database. Together, these components form an intelligent instructional platform capable of natural language interaction, semantic diagnosis, and process tracking. The overall architecture is illustrated in Figure 4.

Figure 4. System architecture of the GPT-based interactive platform for reflective thinking in UAV-based STEM education



The front-end interface was built using the Streamlit framework, supporting multi-turn dialogue, weekly task classification, and corpus export functionalities. During in-class activities, students could input task-related inquiries through the interface. These inputs were routed to a locally deployed Ollama language model, which generated context-aware natural language responses based on conversation history. The semantic generation engine was designed using a hybrid prompt logic, combining static templates with dynamic semantic modules. The static templates were aligned with the course structure and predefined reflection frameworks. The dynamic module, on the other hand, analyzed features of student input, such as the presence of negations, erroneous vocabulary, and sentence length, to adjust response strategies accordingly. This allowed the system to tailor its responses in real time and guide students toward deeper reflection and conceptual clarification.

The task-prompt logic module was built on the Sentence-BERT (all-MiniLM-L6-v2) embedding model, using cosine similarity and a hybrid approach combining rule-based filtering and keyword classification. The module identified the semantic intent of student input. For example, if student queries included terms like “drifting,” “no response,” or “sensor failure,” the system generated diagnostic responses such as “Have you considered the flight environment or PID settings?” to prompt error reasoning and strategic adjustment. All system interactions—including user input, GPT responses, prompt types, and parameter configurations—were stored in real time in a MongoDB database. Data could be retrieved and analyzed by week, semantic level, or anonymized student ID. All entries were encrypted and anonymized using hashed identifiers and stored exclusively on local servers to ensure data security and ethical compliance.

3.4. Procedures

To systematically present the impact of the GPT-based system on students’ reflective processes across different task phases, the instructional timeline was divided into three stages aligned with the course weeks. Each stage is described in relation to reflective thinking levels (Habitual Action, Understanding, Reflection, and Critical Reflection) and corresponding higher-order thinking processes.

During Weeks 1–2, the instructional focus was on module selection and task requirement planning. At this early stage, most students demonstrated an operational mindset driven by surface-level feature associations. For example, a student might choose a module simply because "the camera can take pictures," without considering its functional parameters. Their cognitive engagement was largely confined to the Habitual Action and Understanding stages. In this context, the GPT system acted as a strategic initiator, using counter-questions and semantic restructuring to prompt design clarification and premise evaluation. As shown in Figure 5, typical system responses included prompts such as, "Why did you choose this module? Have you considered its update frequency or data stability?"

These questions encouraged students to move from superficial reasoning toward more structured and logical strategic thinking.

Figure 5. Examples of students' technical inquiries and semantic clarification through GPT interactions during UAV system configuration

During Weeks 3–4, the curriculum entered a more technically challenging phase focused on flight control integration and system testing. At this point, students encountered multi-dimensional problems such as sensor failures and PID parameter anomalies. Reflective activities began to shift from isolated procedural responses toward cross-week integration and causal reasoning, indicating the emergence of Analysis-level thinking. As illustrated in Figure 6, when students asked questions such as, "Why does the drone drift during testing?", the GPT system responded with queries like, "Is your testing environment consistent? Could wind speed or sensor placement be causing interaction effects?" Such responses guided students to construct logical reasoning chains and engage in strategic loop formation, deepening their analytical reflection.

Figure 6. Real-time interaction with the GPT system during UAV sensor integration and coding tasks, supporting reflective thinking and strategy refinement.

In Weeks 5–6, students conducted flight demonstrations and presented their final project outcomes. Reflective attention shifted toward synthesizing learning experiences and evaluating the effectiveness of their design decisions. Student responses began to include metacognitive statements such as, "Although the flight was successful, did the result truly align with the original design goal?" This indicated progression into the Critical Reflection stage and the development of Evaluation-level thinking. At this point, the GPT system adopted the role of a metacognitive inquirer, posing questions like, "Does your final output genuinely address the user needs defined in your original task?" Such prompts encouraged students to reflect on the broader value and purpose of their work. This form of semantic feedback not only strengthened evaluative thinking but also supported students in integrating experiences and knowledge for redesign and optimization, thereby demonstrating higher-order thinking at the Create level.

3.5. Instruments

To comprehensively evaluate the effects of integrating generative AI into UAV-focused engineering courses on students' reflective thinking and HOTS, this study adopted a mixed-methods approach combining both quantitative and qualitative data collection and analysis. On the quantitative side, three standardized instruments with established reliability and validity were employed. First, the STEM Literacy Scale, adapted from Kelley and Knowles (2016), consists of 12 items covering four dimensions: scientific inquiry, technological literacy, engineering design, and mathematical reasoning. It uses a five-point Likert scale and demonstrated a Cronbach's alpha of 0.83. Second, the Reflective Thinking Scale, developed by Chang et al. (2025), contains 12 items categorized into four levels: Habitual Action (HA), Understanding (U), Reflection (R), and Critical Reflection (CR). The wording of the items was adjusted based on expert review in educational psychology, yielding a Cronbach's alpha of 0.88. Third, the HOTS Scale, adapted from Zhou et al. (2023), focuses on the analysis and evaluation dimensions of cognitive performance. It includes 8 items and achieved a Cronbach's alpha of 0.86. All three instruments were administered both before and after the intervention to serve as benchmarks for assessing learning gains and cognitive development. The reliability coefficients met the internal consistency criterion recommended by Nunnally (1978), with α values exceeding the 0.70 threshold.

In terms of qualitative data, two types of textual artifacts were collected to complement the quantitative measures and provide insights into students' semantic transformations and cognitive shifts throughout the learning process. The first set comprised weekly reflective entries (150–200 words) written by students, guided by the GPT system prompts. The second set included interaction logs from the GPT dialogue system, consisting of students' questions, GPT responses, time stamps, and week-task alignment data. Each student query was categorized into one of three functional types: clarification, strategy, or critical reflection. Furthermore, each interaction record was tagged by source type, either "system-prompted" or "student-initiated," to enable further analysis of how generative AI feedback supported higher-level semantic output and reflective transformation. All data were encoded and anonymized following ethical guidelines. Quantitative data were used to measure pre- and post-intervention trends, while qualitative data supported in-depth semantic coding and interpretation of students' reflective trajectories and the impact of AI-guided scaffolding.

4. Results

This chapter presents the quantitative and qualitative results examining how the integration of Generative AI into an interdisciplinary UAV-based engineering course influenced students' performance in STEM literacy, reflective thinking, and HOTS. A mixed-methods approach was employed to address each research question, including statistical testing and semantic-level analysis. To examine whether Generative AI enhanced students' STEM literacy, independent sample t test was performed on the experimental group and the control group before and after the test. As shown in Table 2, students in the experimental group scored significantly higher on the STEM Literacy Scale after the intervention (M = 3.60, SD = 0.34) compared to their pre-test scores (M = 3.15, SD = 0.38), with the difference reaching statistical significance ($t(31) = 4.51$, $p < .001^{***}$). The control group also showed a statistically significant improvement in STEM literacy post-intervention (M = 3.32, SD = 0.25) compared to pre-test scores (M = 3.14, SD = 0.33), though the effect was smaller ($t(31) = 2.25$, $p < .05^{*}$).

Table 2. Pre- and post-test results on STEM literacy scales by group

	Mean(SD)		*t*	*df*	*p*
	Pre-test (n=32)	Post-test (n=32)			
EG STEM Scale	3.15(0.38)	3.60(0.34)	4.51	31	$<.001^{***}$
CG STEM Scale	3.14(0.33)	3.32(0.25)	2.25	31	$< .05^{*}$

Note. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

To evaluate the impact of Generative AI on students' reflective thinking, descriptive statistics and an independent samples t-test were conducted for both groups. Figure 7 illustrates changes in Reflective Thinking scores, while Table 3 details the statistical outcomes. The experimental group exhibited a significant increase in Reflective Thinking scores post-intervention (M = 3.68, SD = 0.47) compared to pre-test scores (M = 3.12, SD = 0.33), with the difference reaching statistical significance ($t(31) = -4.75$, $p < .001^{***}$). In contrast, the control group's scores showed no significant difference before (M = 3.09, SD = 0.93) and after the course (M = 3.21, SD = 0.42) ($t(31) = -1.072$, $p = .292$).

Figure 7. Visualized means and standard deviations of students' reflective thinking scores before and after the Generative AI intervention

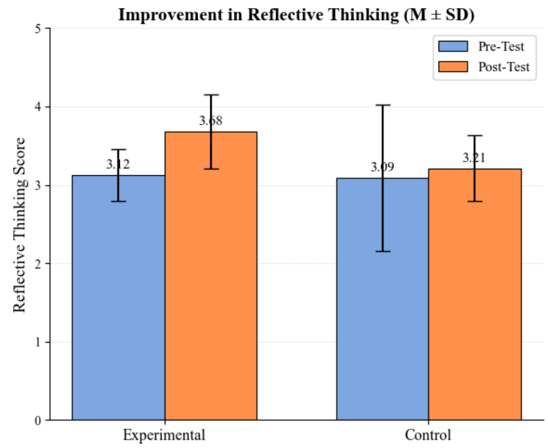


Table 3. Pre- and post-test results on reflective thinking scales by group

	Mean(SD)		t	df	p
	Pre-test (n=32)	Post-test (n=32)			
EG Reflective Thinking	3.12(0.33)	3.68(0.47)	-4.75	31	$<.001^{***}$
CG Reflective Thinking	3.09(0.93)	3.21(0.42)	-1.07	31	.292

Note. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

To evaluate whether generative AI-supported instruction effectively enhanced students' performance in higher-order thinking processes, this study conducted descriptive statistics and independent samples t-tests on the HOTS Scale scores of both the experimental and control groups. The focus was placed on the degree of change and statistical significance across the cognitive dimensions before and after the intervention. Figure 8 illustrates the score changes in the dimensions of higher-order thinking, while Table 4 presents the corresponding statistical test results.

Figure 8. Visualized means and standard deviations of higher-order thinking dimensions across groups

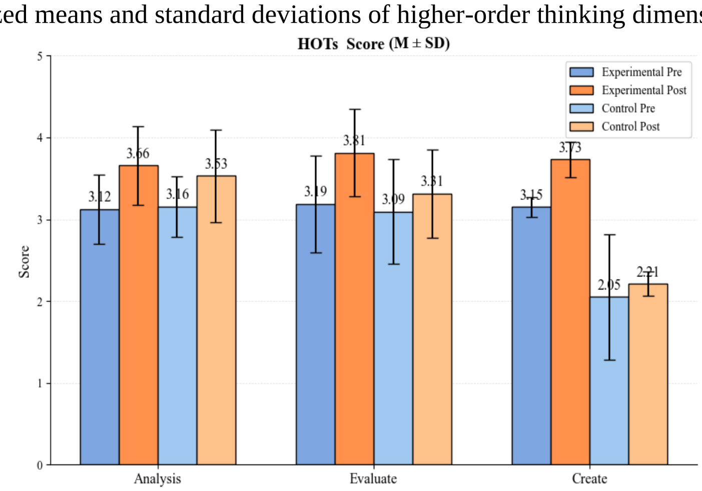


Table 3. Pre- and post-test results on HOTS by group

	Mean(SD)		*t*	*df*	*p*
	Pre-test (n=32)	Post-test (n=32)			
EG Analysis	3.12(0.42)	3.65(0.48)	-4.18	31	$<.001^{***}$
EG Evaluate	3.18(0.59)	0.81(0.53)	-4.45	31	$<.001^{***}$
EG Create	3.15(0.12)	3.73(0.22)	-4.11	31	$<.001^{***}$
CG Analysis	3.15(1.36)	3.53(0.56)	-3.21	31	$<.003^{**}$
CG Evaluate	3.09(0.64)	3.31(0.53)	-1.64	31	.109
CG Create	2.05(0.77)	2.21(0.15)	-3.44	31	.121

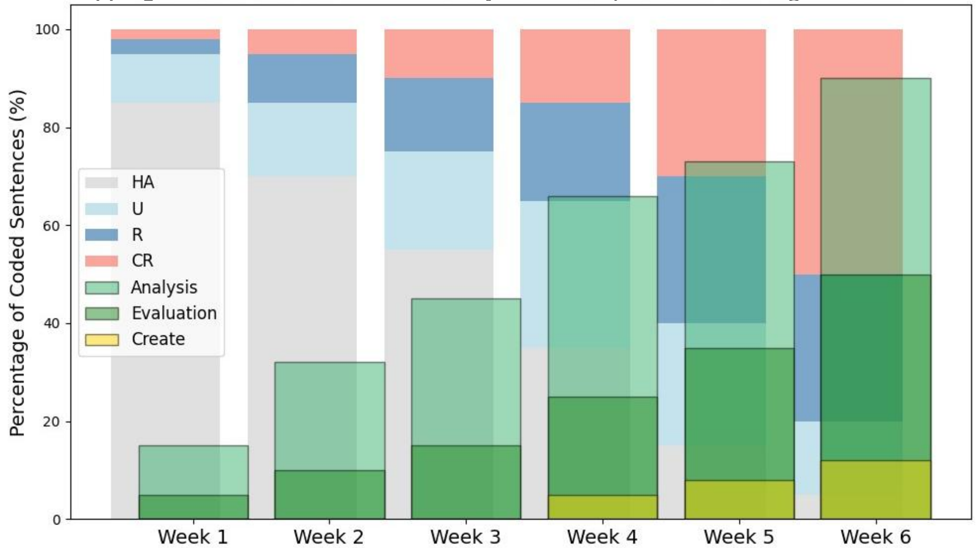
Note. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

In the Analysis dimension, the experimental group's post-test scores (M = 3.65, SD = 0.48) were significantly higher than their pre-test scores (M = 3.12, SD = 0.42), reaching a high level of significance (t = −4.18, $p < .001^{***}$). The control group also showed a statistically significant increase from pre-test (M = 3.15, SD = 1.36) to post-test (M = 3.53, SD = 0.56), (t = −3.21, p = .003), although the larger standard deviation indicated greater score variability within the group. In the Evaluation dimension, the experimental group demonstrated a significant increase from pre-test (M = 3.18, SD = 0.59) to post-test (M = 3.81, SD = 0.53), with the result reaching statistical significance (t = −4.45, $p < .001^{***}$). The control group, however, did not show a statistically significant change in the same dimension (t = −1.64, p = .109). In the Create dimension, the experimental group again showed a statistically significant improvement from pre-test (M = 3.15, SD = 0.12) to post-test (M = 3.73, SD = 0.22), (t = −4.11, $p < .001^{***}$). The control group, on the other hand, exhibited no statistically significant change (t = −3.44, p = .121).

To further investigate the developmental trajectory of students' reflective thinking during the course, a thematic qualitative analysis was conducted on the weekly reflection entries and GPT dialogue transcripts submitted by the experimental group over the six-week period. The analysis followed a three-phase coding process to identify semantic transformations and cognitive progression. In the open coding phase, two researchers independently segmented all textual data into semantic units and labeled them using six thematic categories: (1) strategy-based decision-making, (2) error diagnosis and correction, (3) learning difficulties and uncertainty, (4) clarification of functional needs, (5) outcome evaluation, and (6) value-oriented reflection. Dual coding was employed to reduce individual researcher bias and to enhance the credibility and dependability of the results. To verify inter-coder reliability, Cohen's Kappa was calculated, yielding an average value of 0.86. According to the classification by Landis and Koch (1977), this value reflects an "almost perfect agreement," indicating a high level of consistency in data interpretation.

In the axial coding phase, data were simultaneously coded using two analytical frameworks. One addressed the levels of reflective thinking, including HA, U, R, and CR, as defined by Kember et al. (2008). The second focused on higher-order thinking skills: Analysis, Evaluation, and Create. Each semantic unit was assigned at least one code from each framework to capture multidimensional meaning. The trajectory pattern analysis phase examined weekly shifts in the proportions of coded reflective and cognitive levels to trace the evolution of students' thinking. As shown in Figure 9, during Week 1, over 85% of students' statements fell into the HA category, indicating a predominance of routine actions and surface-level descriptions. By Week 3, the proportions of U and R increased noticeably, while HA dropped below 55%. By Week 6, the CR category had risen to 50%, reflecting a shift from experiential description to more complex evaluations of task logic and value integration.

Figure 9. Overlapping distribution of students' weekly reflective processes and higher-order thinking levels



In terms of higher-order thinking, Analysis-level statements accounted for 75% of entries in Week 1 and gradually increased to 100% by Week 6. Evaluation-level responses rose from 25% in Week 1 to 51% in Week 6. Meanwhile, Create-level statements began to emerge in Week 4 and reached 32% by Week 6, with content focusing on proposing alternative strategies, redesigning workflows, and reconfiguring problem-solving approaches. These changes indicate that students were not only engaged in reflecting on their experiences but also actively constructing novel solutions, demonstrating an increased capacity for innovation and conceptual transformation. The content of the reflections likewise showed a transition from technical analysis toward more strategic reasoning, task justification, and outcome-based evaluation.

5. Discussion

The findings of this study suggest that the implementation of Generative AI within an engineering-oriented UAV project course has a positive impact on students' STEM literacy, HOTS, and reflective thinking development. By integrating the results from both quantitative assessments and qualitative discourse analysis, this research aims to elucidate the cognitive mechanisms through which Generative AI fosters strategic learning adjustments, deepens semantic understanding, and supports the integration of cross-disciplinary competencies. Regarding STEM literacy, both the experimental and control groups demonstrated significant post-intervention gains. However, the experimental group achieved a higher average score and a greater degree of improvement. This indicates that the semantic prompts and interactive feedback generated by Generative AI may contribute additional benefits to the construction of STEM literacy. These results are consistent with the argument proposed by Yeni et al. (2024), who emphasized that literacy development relies on the ability to translate contextual knowledge and make interdisciplinary connections, rather than solely accumulating subject-specific content. Within this framework, Generative AI served as a literacy translator, supporting students in transforming procedural details—such as sensor configurations, module selection, and data frequency—into scientific hypothesis building and engineering strategy formulation. For instance, during Week 3, when students completed flight control integration, they were able to clearly articulate how signal latency might affect data stability. This suggests that AI-generated prompts guided students to extract scientific reasoning and technical logic from their operational experiences. These findings reflect the bridge between technology and knowledge construction described by Tritico and Korach (2024).

In terms of STEM literacy, students moved beyond a simple focus on the successful operation of drones. Instead, they began to consider the underlying logic of assembly, the appropriateness of components for specific tasks, and the prevention of potential failures. Feedback from Generative AI triggered reflective thinking about the rationale for design decisions and optimization strategies, as also discussed by Fajrina et al. (2020) and Abedi et al. (2023). The development of mathematical literacy was particularly evident in situations such as PID parameter adjustments, prediction of model outputs, and correction of data errors. Students demonstrated the ability to interpret graph trends to understand variable influences and modify their strategies accordingly. This reflects the kind of applied mathematical thinking emphasized by Baričević and Luić (2023). Overall, the GPT system guided students in developing structured awareness of different dimensions of STEM literacy. For example, students became more conscious of how data accuracy relates to mathematical literacy, how selecting specific modules reflects technical literacy, and how task decomposition involves engineering literacy. This suggests a notable improvement in cognitive awareness and the ability to recognize literacy structures.

In addition to supporting integrated STEM literacy, Generative AI also significantly enhanced students' performance in the higher-order thinking domains of analysis and evaluation. Students were able to deconstruct information, identify reasoning flaws, and assess the effectiveness and consequences of strategic choices. This indicates a cognitive shift from basic judgment to strategic verification and value-based reasoning, aligning with the observations of Clarke and Konak (2025). The discourse analysis revealed that GPT prompted students to engage in logical deconstruction and strategic reflection through open-ended questions, such as "Have you considered the interaction between wind conditions and PID settings?" These responses served dual functions by promoting multi-causal reasoning and encouraging students to transition from isolated operations to systems-level thinking. This progression corresponds closely with the revised version of Bloom's taxonomy proposed by Anderson and Krathwohl (2001), particularly in the transition from analysis to evaluation. Students also demonstrated advanced verbal reasoning by articulating statements such as "Although this module is simpler, does it compromise data stability?" or "Could insufficient sensor frequency result in delayed decision-making?" These expressions reflect the emergence of student-generated evaluative criteria and value-based comparisons. Such behavior illustrates the characteristics of strategic cognition that were emphasized in the work of Fajrina et al. (2020) and Tran et al. (2025).

In terms of reflective development, students' statements progressively transitioned from HA to U, R, and ultimately to CR. By Week 6, CR-level utterances accounted for 50 percent of the total, indicating a stable formation of deep reflective capabilities. This developmental trajectory not only provides empirical support for the four-level reflection model proposed by Kember et al. (2008), but also aligns with Dewey's (1933) conception of reflection as a process of reconstructing experience. Notably, from Week 3 onward, the GPT system increasingly functioned as a cognitive scaffold. Through semantic reframing and strategic questioning, it guided students to transform operational failures into reflective inquiries. For instance, instead of merely noting a "module failure," students began to formulate follow-up questions prompted by GPT feedback, such as "Could sensor malfunction be caused by sudden changes in lighting?" This behavior exemplifies what Moon et al. (2024) describe as semantic reorganization, a cognitive ability in which learners restructure meaning to generate new understanding.

The study also found that Generative AI not only elevated the level of reflection content but also supported students in developing their own chains of self-questioning. This reflects its potential as a facilitator of metacognitive regulation. Qualitative data revealed that students' post-class reflections exhibited a staged structure, beginning with awareness of a problem, followed by hypothesis verification, and culminating in strategy adjustment. Importantly, these developments were primarily driven by AI-mediated dialogue rather than teacher-led interventions. This finding resonates with the theoretical perspective proposed by Chauhan and Sevda (2023), who described AI interfaces as a semantic mirror that reflects and promotes strategic shifts in learner behavior.

Students also demonstrated notable changes in the linguistic features of their responses. These included progressively longer interaction sequences, increased rhetorical density, and a higher proportion of self-initiated questions. Such changes suggest that learners were increasingly able to externalize their implicit cognitive strategies into explicit language, allowing for more conscious semantic interpretation and logical articulation. This phenomenon supports the argument by Preiksaitis and Rose (2023), who asserted that Generative AI holds the potential to facilitate the verbalization of internal thought processes and enhance students' self-monitoring abilities. This, in turn, enables the construction of structured reflective processes and the translation of meaning. In summary, within a STEM-oriented curriculum, Generative AI has evolved beyond its initial role as a task-solving tool. It has become a multifaceted learning partner that supports students in deepening reflection, integrating strategies, and constructing disciplinary literacy awareness. These findings carry significant implications for both instructional practices and theories of learning.

6. Conclusion

This study explored the potential benefits of Generative AI in supporting students' development of STEM literacy, HOTS, and reflective capacity within an engineering-oriented interdisciplinary curriculum. The findings suggest that when an AI system is designed as a learning partner equipped with language comprehension and strategic prompting capabilities, it can effectively enhance students' integrative performance in scientific inquiry, technological application, engineering design, and mathematical reasoning. Through sustained interaction with the Generative AI system, students gradually exhibited deeper levels of linguistic reflection and strategic judgment. Their verbal expressions shifted from mere operational descriptions to semantically rich articulations involving logical reasoning and evaluative thinking. This progression indicates that Generative AI not only provides real-time linguistic support but also serves as a mediating agent in facilitating reinterpretation of learning tasks and promoting self-regulation. By integrating theoretical frameworks of higher-order thinking and levels of reflective practice, the study demonstrates that Generative AI's feedback extends beyond information delivery to activate cognitive reorganization and semantic deepening during the learning process. From an instructional design perspective, Generative AI should not be regarded as a unidirectional support tool, but rather as a dynamic co-constructive agent in teacher-student knowledge building. To fully harness AI's potential for supporting deep learning, educators should consider the linguistic structure of prompts, timing of interaction, and alignment with task scenarios when designing such systems to ensure meaningful cognitive transformation.

Despite the promising outcomes, certain limitations warrant attention. The study focused on first-year engineering students, whose language expression styles and task receptiveness may affect the generalizability of AI-supported interaction across broader educational contexts. Moreover, the current Generative AI system has limited capacity to detect nuanced semantic deviations and contextual cues, which may lead to overly technical feedback that overlooks conceptual transformation or creative generation. Future research may examine the interplay between task types and semantic prompting strategies and develop adaptive linguistic models that dynamically adjust the depth of feedback based on student utterance features. In addition, the development of analytic dashboards that track real-time semantic shifts and cognitive trajectories is recommended to improve AI's precision in learning monitoring and individualized diagnostics. In conclusion, this study affirms the multifunctional role of Generative AI in the learning process as a provider of linguistic scaffolding, strategy elicitation, and cognitive facilitation. These insights not only extend the boundaries of digital tool applications in education but also offer a theoretical and empirical foundation for designing future intelligent learning systems and advancing AI-related educational literacy.

References

Abedi, M., Alshybani, I., Shahadat, M. R. B., & Murillo, M. (2023). *Beyond traditional teaching: The potential of large language models and chatbots in graduate engineering education*. Qeios. https://doi.org/10.32388/MD04B0

Abuhassna, H., Van, N. T., Yahaya, N., Zakaria, M. A. Z. M., Awae, F., Al Zitawi, D. U. D., & Bayoumi, K. (2022). Strategies for Successful Blended Learning-A Bibliometric Analysis and Reviews. *Int. J. Interact. Mob. Technol.*, *16*(13), 66-80. https://doi.org/10.3991/ijim.v16i13.30739

Ahmed, Z., Ahmed, M., Ahmad, M., Asif, M., & Ameen, M. (2024). The generative AI landscape in education: Mapping the terrain of opportunities, challenges, and student perception. *IEEE Access, 12,* 147023–147050. https://doi.org/10.1109/ACCESS.2024.3461874

AlAli, R. (2024). Enhancing 21st century skills through integrated STEM education using project-oriented problem-based learning. *Geo Journal of Tourism and Geosites, 53*(2), 421–430. https://doi.org/10.30892/gtg.53205-1217

Aldalur, I., & Sagarna, X. (2023). Improving programming learning in engineering students through discovery learning. *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje*, *18*(3), 239-249. https://doi.org/10.1109/RITA.2023.3301409

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives: complete edition*. Addison Wesley Longman, Inc..

Baričević, M., & Luić, L. (2023). From active learning to innovative thinking: The influence of learning the design thinking process among students. *Education sciences*, *13*(5), 455. https://doi.org/10.3390/educsci13050455

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1964). *Taxonomy of educational objectives*(Vol. 2). New York: Longmans, Green.

Castro, V. D. S., & Oliveira, S. R. B. (2023). Diversity in software design and construction teaching: A systematic literature review. *Education Sciences, 13*(3), 303. https://doi.org/10.3390/educsci13030303

Çetinkaya, M., Mercan, G., & Varol Selçuk, Z. (2025). Mapping the evolution of systems thinking in science education: A bibliometric analysis (January 2010 – December 2024). *International Journal of Education Technology and Scientific Researches*, *10*(30), 55–74. http://dx.doi.org/10.35826/ijetsar.767

Chang, C.-Y., Lin, H.-C., Yin, C., & Yang, K.-H. (2025). Generative AI-assisted reflective writing for improving students' higher order thinking: Evidence from quantitative and epistemic network analysis. *Educational Technology & Society*, *28*(1), 270–285. https://www.jstor.org/stable/48810719

Chaudhry, I. S., Sarwary, S. A. M., El Refae, G. A., & Chabchoub, H. (2023). Time to revisit existing student's performance evaluation approach in higher education sector in a new era of ChatGPT—A case study. *Cogent Education*, *10*(1), 2210461. https://doi.org/10.1080/2331186X.2023.2210461

Chauhan, G., & Sevda, S. (2023). *Solid Waste Management*. CRC Press.

Chen, B., Zhang, Z., Langrené, N., Zhu, S. (2025). Unleashing the potential of prompt engineering for large language models. *Patterns*. *6*(6), 101260. https://doi.org/10.1016/j.patter.2025.101260

Chen, H. J., & Lin, Y. F. (2025). Exploring the impact of different tools on metacognitive performance. In M. Kurosu & A. Hashizume (Eds.), *Human-Computer Interaction. HCII 2025. Lecture Notes in Computer Science* (Vol. 15767, pp. 131–149). Springer. https://doi.org/10.1007/978-3-031-93838-2_9

Chen, X., Zou, D., Xie, H., & Wang, F. L. (2023). Metaverse in education: Contributors, cooperations, and research themes. *IEEE Transactions on Learning Technologies*, *16*(6), 1111-1129. https://doi.org/10.1109/TLT.2023.3277952

Clarke, C. J. S. F., & Konak, A. (2025). The impact of AI use in programming courses on critical thinking skills. *Journal of Cybersecurity Education, Research and Practice, 2025*(1), Article 5. https://doi.org/10.62915/2472-2707.1220

Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process*. Heath & Co.

Diez, Y., Kentsch, S., Fukuda, M., Caceres, M. L. L., Moritake, K., & Cabezas, M. (2021). Deep learning in forestry using uav-acquired rgb data: A practical review. *Remote Sensing*, *13*(14), 2837. https://doi.org/10.3390/rs13142837

Doğan, E., & Şahin, F. (2024). Advances in artificial intelligence in education: Leading contributors, current hot topics, and emerging trends. *Participatory Educational Research, 11*(H. Ferhan Odabaşı Gift Issue), 95–113. https://doi.org/10.17275/per.24.96.11.6

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., ... Williams, M. D. (2021). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management, 57*, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Fajrina, S., Lufri, L., & Ahda, Y. (2020). Science, technology, engineering, and mathematics (STEM) as a learning approach to improve 21st century skills: a review. *International Journal of Online & Biomedical Engineering*, *16*(7). https://doi.org/10.3991/ijoe.v16i07.14101

Fan, G., Liu, D., Zhang, R., Pan, L. (2025). The impact of AI-assisted pair programming on student motivation, programming anxiety, collaborative learning, and programming performance: A comparative study with traditional pair programming and individual approaches. *International Journal of STEM Education, 12*(1), Article 16. https://doi.org/10.1186/s40594-025-00537-3

Fang, J. W., Chen, J., Weng, Q. L., Tu, Y. F., Hwang, G. J., & Xia, Y. C. (2025). Effects of a GenAI-based debugging approach integrating the reflective strategy on senior high school students' learning performance and computational thinking. *Educational Technology & Society*, *28*(3), 66-81. https://doi.org/10.30191/ETS.202507_28(3).SP06

Fu, W., Zhang, J., Zhang, D., Li, T., Lan, M., & Liu, N. (2025). An Empirical Study of Adaptive Feedback to Enhance Cognitive Ability in Programming Learning among College Students: A Perspective Based on Multimodal Data Analysis. *Journal of Educational Computing Research*, *63*(3), 532-564. https://doi.org/10.1177/07356331241313126

Garcia, M. B. (2025). Teaching and learning computer programming using ChatGPT: A rapid review of literature amid the rise of generative AI technologies. *Education and Information Technologies.* Advance online publication. https://doi.org/10.1007/s10639-025-13452-5

Gong, X., Li, Z., & Qiao, A. (2025). Impact of generative AI dialogic feedback on different stages of programming problem solving. *Education and Information Technologies, 30*, 9689–9709. https://doi.org/10.1007/s10639-024-13173-1

Gueye, M. L., & Exposito, E. (2024). Hybrid course instructional design for better practice of learning analytics. *Interactive Learning Environments*, *33*(1), 420–439. https://doi.org/10.1080/10494820.2024.2350644

Guo, Q., Zhen, J., Wu, F., He, Y., & Qiao, C. (2025). Can Students Make STEM Progress With the Large Language Models (LLMs)? An Empirical Study of LLMs Integration Within Middle School Science and Engineering Practice. *Journal of Educational Computing Research*, *63*(2), 372-405. https://doi.org/10.1177/07356331241312365

Hazarika, B., Huang, Y., Ahmed, A., Rana, N., Li, X., Li, S., & Gupta, R. (2025). Generative AI-augmented graph reinforcement learning for adaptive UAV swarm optimization. *IEEE Internet of Things Journal*, *12*(8), 9508–9524. https://doi.org/10.1109/JIOT.2025.3529904

Holzinger, A., Saranti, A., Angerschmid, A., Retzlaff, C. O., Gronauer, A., Pejakovic, V., Medel-Jimenez, F., Krexner, T., Gollob, C., & Stampfer, K. (2022). Digital Transformation in Smart Farm and Forest Operations Needs Human-Centered AI: Challenges and Future Directions. *Sensors*, *22*(8), 3043. https://doi.org/10.3390/s22083043

Jauhari, S. (2024). How Do Big Data and Generative AI Dawn on Computational Biology?. In *Unraveling New Frontiers and Advances in Bioinformatics* (pp. 193-228). Singapore: Springer Nature Singapore.

Jiang, M. Y. C., Jong, M. S. Y., Chai, C. S., Huang, B., Chen, G., Lo, C. K., & Wong, F. K. K. (2024). They believe students can fly: A scoping review on the utilization of drones in educational settings. *Computers & Education*, *205*, 105113. https://doi.org/10.1016/j.compedu.2024.105113

Jiao, J., Afroogh, S., Chen, K., Atkinson, D., & Dhurandhar, A. (2024). The global landscape of academic guidelines for generative AI and large language models. *arXiv Preprint arXiv:2406.18842*. https://doi.org/10.48550/arXiv.2406.18842

Kaleem, Z., Orakzai, F. A., Ishaq, W., Latif, K., Zhao, J., & Jamalipour, A. (2024). Emerging trends in UAVs: From placement, semantic communications to generative AI for mission-critical networks. *IEEE Transactions on Consumer Electronics*. https://doi.org/10.1109/TCE.2024.3434971

Kelley, T. R., & Knowles, J. G. (2016). A conceptual framework for integrated STEM education. *International Journal of STEM Education*, *3*(1), 11. https://doi.org/10.1186/s40594-016-0046-z

Kember, D., McKay, J., Sinclair, K., & Wong, F. K. Y. (2008). A four-category scheme for coding and assessing the level of reflection in written work. *Assessment & Evaluation in Higher Education*, *33*(4), 369–379. https://doi.org/10.1080/02602930701293355

Kim, P., Wang, W., & Bonk, C. J. (2025). Generative AI as a Coach to Help Students Enhance Proficiency in Question Formulation. *Journal of Educational Computing Research*, *63*(3), 565-586. https://doi.org/10.1177/07356331251314222

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

Lang, Q., Wang, M., Yin, M., Liang, S., & Song, W. (2025). *Transforming education with generative AI (GAI): Key insights and future prospects*. *IEEE Transactions on Learning Technologies*, *18*, 230–242. https://doi.org/10.1109/TLT.2025.3537618

Lee, Y. F., Lee, L. S., & Nguyen, H. B. N. (2023). A comparison of STEM education status and trends in five highly competitive countries in the Asia-Pacific region. *Journal of Asian Vocational Education and Training*, *16*, 2.

Lievens, F., Harrison, S. H., Mussel, P., & Litman, J. A. (2022). Killing the cat? A review of curiosity at work. *Academy of Management Annals*, *16*(1), 179-216. https://doi.org/10.5465/annals.2020.0203

Lin, H. C. K., Feng, C. H., & Liu, Y. L. E. (2025). Generative AI-enhanced UX storyboarding: Transforming novice designers' engagement and reflective thinking for IoMT product design. *International Journal of Technology and Design Education*. https://doi.org/10.1007/s10798-025-09993-y

Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2023). Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM Computing Surveys*, *55*(9), Article 195. https://doi.org/10.1145/3560815

Meneses, E. (2023). Impact of the implementation of ChatGPT in education: A systematic review. *Computers*, *12*(8), 153. https://doi.org/10.3390/computers12080153

Meng, X. P., Guo, X. G., Fang, J. W., Chen, J., & Huang, L. F. (2025). Fostering pre-service teachers' generative AI literacy and critical thinking: An RSCQA approach. *Educational Technology & Society*, *28*(3), 202-225. https://doi.org/10.30191/ETS.202507_28(3).TP01

Mezirow, J. (1991). *Transformative dimensions of adult learning*. Jossey-Bass, 350 Sansome Street, San Francisco, CA 94104-1310.

Min, T., Lee, B., & Jho, H. (2025). Integrating generative artificial intelligence in the design of scientific inquiry for middle school students. *Education and Information Technologies*, 1-32. https://doi.org/10.1007/s10639-025-13410-1

Moon, J., Yeo, S., Si, Q., & Ijeluola, A. S. (2024). A scoping review of game-based learning for mathematics teacher education. *International Journal of Mathematical Education in Science and Technology*, 1-29. https://doi.org/10.1080/0020739X.2024.2337934

Namaziandost, E., Heydarnejad, T., & Azizi, Z. (2023). The impacts of reflective teaching and emotion regulation on work engagement: Into prospect of effective teaching in higher education. *Teaching English Language*, *17*(1), 139-170. https://doi.org/10.22132/tel.2022.164264

Nunnally, J. C. (1978). *Psychometric Theory: 2d Ed*. McGraw-Hill.

Preiksaitis, C., & Rose, C. (2023). Opportunities, challenges, and future directions of generative artificial intelligence in medical education: scoping review. *JMIR medical education*, *9*, e48785. https://doi.org/10.2196/48785

Rasul, T., Nair, S., Kalendra, D., Robin, M., Santini, F. de O., Ladeira, W. J., Sun, M., Day, I., Rather, R. A., & Heathcote, L. (2023). The role of ChatGPT in higher education: Benefits, challenges, and future research directions. *Journal of Applied Learning and Teaching*, *6*(1), 41–56. https://doi.org/10.37074/jalt.2023.6.1.29

Reddy, G. S., Rizvi, H. A., Kochhar, H. K., Shaikh, A., Udhayakumar, R., & Dhanraj, J. A. (2024, September). Exploring the Role of Natural Language Processing in Educational Chatbots: A Machine Learning Approach. In *2024 7th International Conference on Contemporary Computing and Informatics (IC3I)* (Vol. 7, pp. 441-445). IEEE. https://doi.org/10.1109/IC3I61595.2024.10828612

Runnel, P., Aljas, A., & Jaanits, J. (2024). Co-creation of knowledge and learning in participatory processes: Case study of the DIY exhibitions from the Estonian National Museum. *The International Journal of the Inclusive Museum*, *18*(1), 43–61. https://doi.org/10.18848/1835-2014/CGP/v18i01/43-61

Sandhaus, H., Gu, Q., Parreira, M. T., & Ju, W. (2024). Student Reflections on Self-Initiated GenAI Use in HCI Education. *arXiv preprint arXiv:2410.14048*. https://doi.org/10.48550/arXiv.2405.01467

Sapkota, R., Roumeliotis, K. I., & Karkee, M. (2025). UAVs Meet Agentic AI: A Multidomain Survey of Autonomous Aerial Intelligence and Agentic UAVs. *arXiv preprint arXiv:2506.08045*. https://doi.org/10.48550/arXiv.2506.08045

Sarker, S., Susarla, A., Gopal, R., & Thatcher, J. B. (2024). Democratizing knowledge creation through human-AI collaboration in academic peer review. *Journal of the Association for Information Systems, 25*(1), 158–171. https://doi.org/10.17705/1jais.00872

Satar, M. S., Alharthi, S., Omeish, F., Alshibani, S. M., & Saqib, N. (2024). Digital Learning Orientation and Entrepreneurial Competencies in Graduates: Is Blended Learning Sustainable?. *Sustainability*, *16*(17), 7794. https://doi.org/10.3390/su16177794

Schön, D. A. (1987). *Educating the reflective practitioner.* Jossey-Bass.

Shadiev, R., & Yi, S. (2023). A systematic review of UAV applications to education. *Interactive Learning Environments*, *31*(10), 6165-6194. https://doi.org/10.1080/10494820.2022.2028858

Shehata, A. O., Noroozinejad Farsangi, E., Mirjalili, S., & Yang, T. Y. (2024). A State-of-the-Art Review and Bibliometric Analysis on the Smart Preservation of Heritages. *Buildings*, *14*(12), 3818. https://doi.org/10.3390/buildings14123818

Shi, J., Jain, R., Doh, H., Suzuki, R., & Ramani, K. (2023). An HCI-centric survey and taxonomy of human-generative-AI interactions. *arXiv preprint arXiv:2310.07127*. https://doi.org/10.48550/arXiv.2310.07127

Subramaniam, R. C., Morphew, J. W., Rebello, C. M., & Rebello, N. S. (2025). *Presenting a STEM ways of thinking framework for engineering design-based physics problems. Physical Review Physics Education Research, 21*(1), 010122. https://doi.org/10.1103/PhysRevPhysEducRes.21.010122

Sukackė, V., Pereira de Carvalho Guerra, A. O., Ellinger, D., Carlos, V., Petronienė, S., Gaižiūnienė, L., Blanch, S., Marbà-Tallada, A., & Brose, A. (2022). Towards active evidence-based learning in engineering education: A systematic literature review of PBL, PjBL, and CBL. *Sustainability*, *14*(21), 13955. https://doi.org/10.3390/su142113955

Sun, G., Xie, W., Niyato, D., Du, H., Kang, J., Wu, J., Sun, S., & Zhang, P. (2024). *Generative AI for advanced UAV networking*. *IEEE Network*. https://doi.org/10.1109/MNET.2024.3494862

Thararattanasuwan, K., & Prachagool, V. (2024). Exploring Perspectives of Teacher Students toward Generative AI Technologies. *International Education Studies*, *17*(5), 22-28. https://doi.org/10.5539/ies.v17n5p22

Tran, C., James, B., Allen, V., de Castro, R. O., & Sanin, C. (2025). Using Generative Artificial Intelligence in learning and teaching: An empirical analysis on academic staff's perspectives. *Journal of Applied Learning and Teaching*, *8*(1). https://doi.org/10.37074/jalt.2025.8.1.23

Tritico, H. M., & Korach, C. S. (2024, June). Global Engineering Modules that Teach Currency Exchange and International Trade. In *2024 ASEE Annual Conference & Exposition*, Portland, Oregon. https://doi.org/10.18260/1-2--47509

Tutticci, N., Ryan, M., Coyer, F., & Lewis, P. A. (2018). Collaborative facilitation of debrief after high-fidelity simulation and its implications for reflective thinking: student experiences. *Studies in Higher Education*, *43*(9), 1654-1667. https://doi.org/10.1080/03075079.2017.1281238

Wang, W. S., Lin, C. J., Lee, H. Y., Huang, Y. M., & Wu, T. T. (2025). Integrating feedback mechanisms and ChatGPT for VR-based experiential learning: Impacts on reflective thinking and Aiot physical hands-on tasks. *Interactive Learning Environments*, *33*(2), 1770-1787. https://doi.org/10.1080/10494820.2024.2375644

Xue, J., Yang, P., Li, Q., Song, Y., Gelder, P. V., Papadimitriou, E., & Hu, H. (2025). Machine Learning in Maritime Safety for Autonomous Shipping: A Bibliometric Review and Future Trends. *Journal of Marine Science and Engineering*, *13*(4), 746. https://doi.org/10.3390/jmse13040746

Yang, Y., Luo, J., Yang, M., Yang, R., & Chen, J. (2024). From surface to deep learning approaches with Generative AI in higher education: an analytical framework of student agency. *Studies in Higher Education*, *49*(5), 817-830. https://doi.org/10.1080/03075079.2024.2327003

Yeni, S., Grgurina, N., Saeli, M., Hermans, F., Tolboom, J., & Barendsen, E. (2024). Interdisciplinary integration of computational thinking in K-12 education: A systematic review. *Informatics in Education*, *23*(1), 223-278. https://doi.org/10.15388/infedu.2024.08

Yeung, R. C. Y., Sun, D., & Yeung, C. H. (2025). Integrating drone technology in STEM education: Curriculum, pedagogy and learning outcomes. *Education and Information Technologies*, 1-36. https://doi.org/10.1007/s10639-025-13368-0

Yeung, R. C. Y., Yeung, C. H., Sun, D., & Looi, C. K. (2024). A systematic review of Drone integrated STEM education at secondary schools (2005–2023): Trends, pedagogies, and learning outcomes. *Computers & Education*, *212*, 104999. https://doi.org/10.1016/j.compedu.2024.104999

Yilmaz, R., & Karaoglan Yilmaz, F. G. (2023). The effect of generative artificial intelligence (AI)-based tool use on students' computational thinking skills, programming self-efficacy and motivation. *Computers and Education: Artificial Intelligence*, *4*, 100147. https://doi.org/10.1016/j.caeai.2023.100147

Yuan, B., & Hu, J. (2024). Generative AI as a Tool for Enhancing Reflective Learning in Students. *arXiv preprint arXiv:2412.02603*. https://doi.org/10.48550/arXiv.2412.02603

Zaim, M., Arsyad, S., Waluyo, B., Ardi, H., Al Hafizh, M., Zakiyah, M., Syafitri, W., Nusi, A., & Hardiah, M. (2025). Generative AI as a cognitive co-pilot in English language learning in higher education. *Education Sciences*, *15*(6), 686. https://doi.org/10.3390/educsci15060686

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–where are the educators?. *International journal of educational technology in higher education*, *16*, 39. https://doi.org/10.1186/s41239-019-0171-0

Zhou, L., Zhang, H., Xu, Y., Chen, J., Li, X., Wang, Y., & Liu, F. (2025). *Cooperative generative AI for UAV-based scenarios: An intelligent cooperative framework. IEEE Vehicular Technology Magazine*, *20*(2), 44–52. https://doi.org/10.1109/MVT.2025.3544817

Zhou, Y., Gan, L., Chen, J., Wijaya, T. T., & Li, Y. (2023). Development and validation of a higher-order thinking skills assessment scale for pre-service teachers. *Thinking Skills and Creativity*, *48*, 101272. https://doi.org/10.1016/j.tsc.2023.101272