Effects of integrating image-GenAI with game-based learning and argumentation-based instruction on students' learning effectiveness, computational thinking skills, and metacognitive awareness

Ying-Chieh LIN[a] & Ting-Chia HSU[a*]
[a]*National Taiwan Normal University, Taiwan*
*ckhsu@ntnu.edu.tw

Abstract: This study investigates a novel approach integrating image-based generative AI (GenAI), game-based learning (GBL), and argumentation-based instruction to enhance junior high school students' AI learning. Twenty-six students completed a two-day AI program combining lectures, hands-on activities, and a computational thinking (CT) board game with structured argumentation support. Results revealed significant gains in learning effectiveness, CT skills, and GenAI-related metacognition, demonstrating the approach's potential to foster higher-order thinking and computational competencies in AI education. As an exploratory study, the findings provide initial evidence and lay the groundwork for future large-scale investigations.

Keywords: GenAI, argumentation, game-based learning, computational thinking skills, metacognitive awareness

1. Introduction

Artificial Intelligence (AI) is increasingly embedded in education, supporting teaching, curriculum design, learning analytics, and personalized learning pathways (Zhai et al., 2021). Among its branches, Generative AI (GenAI) comprises text- and image-based tools (Lee et al., 2023). While most research has focused on text-based GenAI applications, such as ChatGPT in English writing courses to improve engagement and learning (Xiao et al., 2025), studies on image-based GenAI in classroom instruction remain scarce for enhancing computational thinking learning or program designing. This gap is particularly evident in computer science education, where abstract and complex concepts present persistent learning barriers. To address this challenge, the present study integrates image-GenAI with argumentation-based instruction, a pedagogical strategy that provides structured reasoning support and cultivates critical thinking (Zheng et al., 2023). Although extensively applied in the social sciences (e.g., Romero Ariza et al., 2021), argumentation is rarely implemented in computer science education. This study employs Toulmin's argumentation model (Rismanto et al., 2021) to scaffold reasoning and deepen conceptual understanding. It examines three key constructs: learning effectiveness as cognitive gains (Ji et al., 2022); computational thinking (CT), the ability to systematically analyze problems and design innovative solutions (Kong, 2019); and metacognitive awareness, the capacity to monitor and regulate cognition, especially in GenAI-supported learning (Fleur et al., 2021). By situating Toulmin-based argumentation within an image-GenAI–enhanced learning environment, this study aims to advance students' mastery of abstract concepts, CT skills, and metacognitive performance—offering a novel instructional model for AI-integrated computer science education. This study is exploratory in nature, focusing on examining the feasibility and potential benefits of integrating image-based GenAI, game-based learning, and argumentation-based instruction in computer science education.

2. Literature review

2.1 Generative AI (GenAI) in education

Generative AI (GenAI) employs deep learning models to produce human-like text and images in response to prompts (Lim et al., 2023) and has demonstrated strong potential to enhance learning effectiveness and student performance (Kumar et al., 2023). Text-based tools such as ChatGPT support teachers in lesson design and feedback, as well as assist students in diverse learning tasks, while image-based tools like Midjourney have been shown to improve visual design and creativity in architectural education (Xu & Huang, 2024). However, GenAI also raises ethical concerns, including plagiarism and academic misconduct (Cotton et al., 2023), and its benefits and risks remain contested, underscoring the need for further empirical and practical investigation (Wang et al., 2024).

2.2 Argumentation-based instruction： Toulmin Model

The Toulmin Argument Model, comprising claim, data, warrant, backing, qualifier, and rebuttal, is regarded as the most comprehensive framework for analyzing argument structure (Rismanto et al., 2021) and has been widely applied in cross-disciplinary instruction. In education, structured argumentation fosters deep understanding, multi-perspective thinking, and improved reasoning quality (Nussbaum, 2021). Recent advances in AI have spurred interest in integrating AI with argumentation, with studies showing AI can act as a virtual learning partner, provide feedback in inquiry or writing, and, when combined with gamified learning, boost engagement and perspective diversity (Lin & Hung, 2025). However, despite its analytical strengths, the Toulmin model lacks standardized criteria for evaluating argument quality, posing challenges for novice learners (Nussbaum, 2021).

2.3 Game-based learning (GBL)

Game-Based Learning (GBL) integrates complete games or game mechanisms into teaching, using clear goals, interactive contexts, immediate feedback, and challenging tasks to achieve learning objectives (Ding et al., 2024). Unlike gamification, which adds game elements, GBL positions the game itself as the primary learning medium. Research shows GBL significantly improves learning outcomes and motivation across disciplines, enhancing mathematics performance and persistence (Vankúš, 2021), boosting vocabulary acquisition (Xu et al., 2023), and strengthening computational thinking in computer education (Wang et al., 2023). However, while traditional GBL is well studied, integrating Generative AI (GenAI) to further enhance GBL's impact in computer education remains underexplored.

3. Methodology

3.1 Instructional Materials

This study employed the newly developed AI 2 Robot City board game, created by Professor Ting-Chia Hsu's team from National Taiwan Normal University. The educational board game integrates image recognition and Internet of Things (IoT) technologies to help students effectively enhance their computational thinking abilities through GBL.

On the software side, the MIT App Inventor platform developed by the Massachusetts Institute of Technology was used to design the mobile application. This was combined with the Personal Image Classifier (PIC) web platform for implementing image recognition functionalities. After training a custom model on the PIC platform, users can import the trained model into the App Inventor environment. The model can then be used to control the robot car based on image input. The board game is played in a mission-based format, with two students per group and two groups competing against each other. Players earn points by collecting building materials shown on the mission cards, with each card offering different point values. This encourages students to collaborate and strategize on how to maximize their score within a limited time. The group with the highest score at the end of the game is declared the winner. A photo of A visual representation of the software implementation is shown in Figure 1.

Figure 1. MIT APP Inventor Interface

3.2 Teaching strategy：Argumentation-based instruction

This argumentation framework is designed with reference to the Toulmin Argument Model. Taking the “image recognition model modification” worksheet from the course as an example, students write based on their stated claim, such as: “Therefore, I believe the model needs to be modified a second time.”

The completion process starts from the far left: First is the Grounds, which provide specific facts or evidence supporting the claim, where students indicate whether the model needs to be modified; next is the Qualifier, which specifies the scope or degree of certainty of the claim and explains why the modification would affect AI judgment; then comes the Warrant, which elaborates on the logical relationship between the grounds and the claim, including results or examples that support the effectiveness of the modification; following that is the Backing, where students search online for information to support their reasoning and record the sources; finally, the Rebuttal prompts students to think about and describe any limitations the modification might have on its effectiveness, as shown in Figure 2.

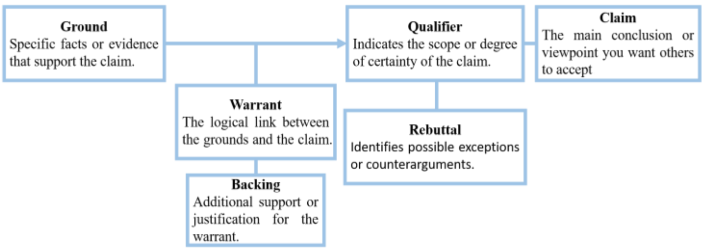

Figure 2. Toulmin Argument Model

3.3 Experimental Procedure

This study investigated the impact of integrating image-based Generative AI (GenAI), argumentation-based instruction, and game-based learning (GBL) on junior high school students’ learning effectiveness, computational thinking (CT), and GenAI-related metacognitive awareness. Twenty-six students (aged 12–15) from northern Taiwan participated in a two-day, 6.5-hours/day learning camp designed to scaffold understanding of abstract AI concepts while fostering cognitive and metacognitive skills. Overall, argumentation-based instruction can be viewed as a pedagogical wrapper that frames the entire learning process, guiding students from task execution toward deeper reflection, reasoning, and collaboration.

On Day 1, students completed a pre-test, studied foundational AI and deep learning concepts, and used the Toulmin argumentation model to visualize reasoning. They generated images via Leonardo AI, translated prompts with Google Translate, and manually redrew outputs to address AI ethics. Students then used the PIC platform to create datasets, train image recognition models, and export them.

On Day 2, students engaged in block-based programming with MIT App Inventor to integrate their models with robotic systems, refined their models following a mini-lecture on optimization, and completed Argumentation Worksheet 2 for reflective reasoning. The program concluded with a custom board game in which teams applied their models to complete mission-based tasks, reinforcing CT, problem-solving, and collaboration. The experimental procedure is shown in Figure 3.

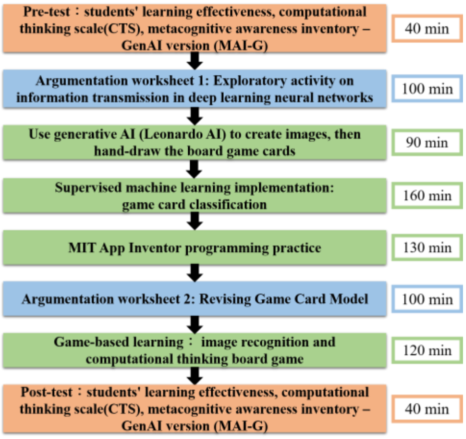

Figure 3. Study Process

Learning effectiveness was measured using a 20-item test (17 multiple-choice, 3 matching; score range 0–100). CT was assessed via Tsai et al.'s (2020) five-dimension Computational Thinking Scale (CTS) which are abstraction, decomposition, algorithmic thinking, evaluation, and generalization. The Cronbach's alpha values for each dimension were .81, .74, .77, .83, and .75, respectively. Metacognitive awareness was evaluated using a Metacognitive Awareness Inventory – Generative AI version (MAI-G) (Moxon, 2023) with an AI ethics dimension from Carolus et al. (2023). The MAI-G consists of three dimensions: cognitive knowledge, regulation knowledge, and AI ethics. The Cronbach's alpha values for each dimension were .79, .77, and .75, respectively.

4. Results

4.1 Student Learning Effectiveness

A comparison of students' performance showed significant improvement from the pretest to the posttest. This indicates that combining argumentation-based instruction with AI-based image generation in game-based learning can significantly enhance students' learning effectiveness. Table 1 presents the summary results of the paired-sample t-test for students' pre- and posttest progress scores.

Table 1 *Summary table of paired t-test for student' learning effectiveness*

	N	Mean	*SD*	*t*	*df*
Pretest	26	70.39	15.70	3.87^{**}	25
Posttest	26	79.00	12.43		

$^{**}p < .01$

4.2 Computational Thinking Scale (CTS)

The CTS results showed significant improvement from the pretest to the posttest. This indicates that combining argumentation-based instruction with AI-based image generation in game-based learning can significantly enhance students' computational thinking. Table 2 presents the summary results of the paired-sample t-test for students' pre- and posttest progress scores.

Table 2 *Summary table of paired t-test for Computational Thinking Scale (CTS)*

	N	Mean	*SD*	*t*	*df*
Pretest	26	3.69	0.69	3.36^{**}	25
Posttest	26	4.13	0.76		

$^{**}p < .01$

4.3 Metacognitive Awareness Inventory – Generative AI version (MAI-G)

The MAI-G results indicated that students' posttest performance showed significant improvement compared to the pretest. This suggests that integrating argumentation-based instruction with game-based learning using image-generative AI can effectively enhance students' metacognitive abilities when using GenAI. Table 3 presents the summary results of the paired-sample t-test for students' progress scores from the pretest to the posttest.

Table 3 *Summary table of paired t-test for MAI-G*

	N	Mean	*SD*	*t*	*df*
Pretest	26	3.92	0.67	4.69^{***}	25
Posttest	26	4.13	0.81		

$^{***}p < .001$

5. Discussion and Conclusion

This study found that integrating argumentation-based instruction with generative AI (GenAI) in game-based learning (GBL) significantly improved students' learning effectiveness, computational thinking, and GenAI-related metacognitive awareness. Structured reasoning combined with interactive, game-oriented activities deepened AI concept understanding, strengthened critical thinking and problem-solving, and built confidence in engaging with AI technologies. The approach fostered active, collaborative learning, encouraging exploration, creativity, and autonomy while reducing apprehension toward complex topics.

The integration of image-based GenAI, game-based learning, and argumentation provides a seamless transition between abstract concept acquisition, interactive gameplay, and reflective reasoning, thereby aligning with the principles of seamless learning by bridging formal instruction with exploratory, self-directed engagement.

Structured argumentation functions as a pedagogical wrapper, enabling students to move beyond task execution toward deeper reflection, reasoning, and collaboration. Integrating argumentation-based instruction with generative AI in game-based learning demonstrates a promising pedagogical model that can simultaneously foster technical competence and higher-order thinking. Educators may flexibly apply this approach across different disciplines to lower cognitive barriers and enhance engagement. Nevertheless, this study has two notable limitations. First, the small sample size (n = 26) constrains the reliability of the findings; expanding the sample would strengthen the robustness of the results. Second, the study adopted a single-group experimental design because it is exploratory in nature, with the primary aim of examining the feasibility and potential benefits of combining image-based generative AI with board-game learning activities. At this stage, research designs typically emphasize preliminary validation and hypothesis generation rather than establishing rigorous causal inferences. Considering the limited sample size and ethical considerations in educational contexts, a single-group pretest–posttest design was employed as a strategy to quickly assess changes in learning outcomes. Future research should include larger and more diverse learner populations, conduct between-group comparisons, explore cross-disciplinary applications, and carry out longitudinal studies to evaluate sustained impacts on motivation, problem-solving, and innovation, thereby providing a more comprehensive understanding of its long-term educational value.

Acknowledgements

This study is partially supported by the National Science and Technology Council under the contract number NSTC 114-2410-H-003-026-MY3.

6. References

Carolus, A., Koch, M. J., Straka, S., Latoschik, M. E., & Wienrich, C. (2023). Mails - meta ai literacy scale: Development and testing of an AI literacy questionnaire based on well-founded competency models and psychological change- and meta-competencies. *Computers in Human Behavior: Artificial Humans*, 1(2), 100014. https://doi.org/10.1016/j.chbah.2023.100014

Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in education and teaching international*, 61(2), 228-239. https://doi.org/10.1080/14703297.2023.2190148

Ding, A. C. E., & Yu, C. H. (2024). Serious game-based learning and learning by making games: Types of game-based pedagogies and student gaming hours impact students' science learning outcomes. *Computers & Education*, 218, 105075. https://doi.org/10.1016/j.compedu.2024.105075

Fleur, D. S., Bredeweg, B., & van den Bos, W. (2021). Metacognition: ideas and insights from neuro- and educational sciences. *npj Science of Learning*, 6(1), 13. https://doi.org/10.1038/s41539-021-00089-5

Ji, M., Luo, Z., Feng, D., Xiang, Y., & Xu, J. (2022). Short- and long-term influences of flipped classroom teaching in physiology course on medical students' learning effectiveness. *Frontiers in Public Health*, 10. https://doi.org/10.3389/fpubh.2022.835810

Kong, S. C. (2019). Components and methods of evaluating computational thinking for fostering creative problem-solvers in senior primary school education. In Computational thinking education (pp. 119-141). *Singapore: Springer Singapore*. https://doi.org/10.1007/978-981-13-6528-7_8

Kumar, T., Kait, R., Ankita, & Malik, A. (2023). The role of Generative Artificial Intelligence (GAI) in education: a detailed review for enhanced learning experiences. In International Conference on Entrepreneurship, Innovation, and Leadership (pp. 195-207). *Singapore: Springer Nature Singapore*. https://doi.org/10.1007/978-981-97-1682-1_17

Lee, U., Han, A., Lee, J., Lee, E., Kim, J., Kim, H., & Lim, C. (2023). Prompt aloud!: Incorporating image-generative AI into steam class with learning analytics using prompt data. *Education and Information Technologies*, 29(8), 9575–9605. https://doi.org/10.1007/s10639-023-12150-4

Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The international journal of management education*, 21(2), 100790.https://doi.org/10.1016/j.ijme.2023.100790

Lin, Y. R., & Hung, C. Y. (2025). The Synergistic Effects in an AI-Supported Online Scientific Argumentation Learning Environment. *Computers & Education*, 105251. https://doi.org/10.1016/j.compedu.2025.105251

Moxon, J. (2023). Psychometric evaluation of abridged versions of the metacognitive awareness inventory in the Japanese population. *Current Psychology*, 42(16), 13228-13230. https://doi.org/10.1007/s12144-022-02763-2

Nussbaum, E. M. (2021). Critical integrative argumentation: Toward complexity in students' thinking. *Educational Psychologist*, 56(1), 1-17. https://doi.org/10.1080/00461520.2020.1845173

Rismanto, R., Asri, A. N., Andoko, B. S., Hirashima, T., & Leonanta, A. A. (2021). A Preliminary Study: Toulmin Arguments in English Reading Comprehension for English as Foreign Language Students. In 2021 International Conference on Electrical and Information Technology (IEIT) (pp. 6-10). *IEEE*. https://doi.10.1109/IEIT53149.2021.9587398

Romero Ariza, M., Quesada Armenteros, A., & Estepa Castro, A. (2024). Promoting critical thinking through mathematics and science teacher education: the case of argumentation and graphs interpretation about climate change. *European Journal of Teacher Education*, 47(1), 41-59. https://doi.org/10.1080/02619768.2021.1961736

Tsai, M.-J., Liang, J.-C., & Hsu, C.-Y. (2020). The Computational Thinking Scale for Computer Literacy Education. *Journal of Educational Computing Research*, 59(4), 579-602. https://doi.org/10.1177/0735633120972356

Vankúš, P. (2021). Influence of game-based learning in mathematics education on students' affective domain: A systematic review. Mathematics, 9(9), 986. https://doi.org/10.3390/math9090986

Wang, N., Wang, X., & Su, Y. S. (2024). Critical analysis of the technological affordances, challenges and future directions of Generative AI in education: a systematic review. *Asia Pacific Journal of Education*, 44(1), 139–155. https://doi.org/10.1080/02188791.2024.2305156

Wang, X., Cheng, M., & Li, X. (2023). Teaching and learning computational thinking through game-based learning: a systematic review. *Journal of Educational Computing Research*, 61(7), 1505-1536. https://doi.org/10.1177/07356331231180951

Xiao, Y., Li, D., & Guo, K. (2025). Using ChatGPT to bring non-player characters to life: Effects on students' storyline-driven game-based writing learning. *Computers & Education*, 105414. https://doi.org/10.1016/j.compedu.2025.105414

Xu, C., & Huang, Y. (2024). Technological Innovation in Architectural Design Education: Empirical Analysis and Future Directions of Midjourney Intelligent Drawing Software. Buildings, 14(10), 3288. https://doi.org/10.3390/buildings14103288

Xu, M., Luo, Y., Zhang, Y., Xia, R., Qian, H., & Zou, X. (2023). Game-based learning in medical education. *Frontiers in public health*, *11*, 1113682. https://doi.org/10.3389/fpubh.2023.1113682

Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., ... & Li, Y. (2021). A Review of Artificial Intelligence (AI) in Education from 2010 to 2020. *Complexity*, 2021(1), 8812542. https://doi.org/10.1155/2021/8812542

Zheng, X. L., Huang, J., Xia, X. H., Hwang, G. J., Tu, Y. F., Huang, Y. P., & Wang, F. (2023). Effects of online whiteboard-based collaborative argumentation scaffolds on group-level cognitive regulations, written argument skills and regulation patterns. *Computers & education*, 207, 104920. https://doi.org/10.1016/j.compedu.2023.104920