Scaffolding AI Literacy Through Student-AI Collaboration in Chatbot Development

Shan Zhang
University of Florida
zhangshan@ufl.edu

Anthony F. Botelho
University of Florida
abotelho@coe.ufl.edu

ABSTRACT

Given the growing integration of artificial intelligence (AI) into education, it is essential to equip students with the knowledge and skills necessary to become creators and designers of AI technologies, rather than passive consumers. A key component of this shift is to develop students' AI literacy, which encompasses not only an understanding of AI concepts but also the ability to apply, evaluate, and innovate with AI in meaningful ways. This study explores how Generative AI (GenAI) can foster AI literacy by providing real-time scaffolding in AMBY (AI Made By You), a conversational chatbot development platform. While chatbot creation provides students with hands-on experience in AI design, including natural language processing and machine learning, the quality of student-created chatbots varies, particularly in the diversity, depth, and effectiveness of their training, response, and testing phrases. To address this, we propose leveraging GenAI-driven support as a collaborator to assist students in generating richer, more contextually diverse, and linguistically varied chatbot responses through adaptive feedback and content refinement suggestions. Moreover, this study aims to develop methods for assessing AI literacy gains by analyzing dialogue interactions and log data, mapping them to established AI literacy competency frameworks. Beyond improving chatbot quality, this research also examines how GenAI-enhanced scaffolding enhances AI literacy while increasing conceptual understanding, problem-solving skills, and self-efficacy across disciplines. The findings will contribute to the design of AI-enhanced learning environments, demonstrating how GenAI can support students in developing high-quality chatbots while deepening their AI literacy and disciplinary knowledge.

Keywords

AI Literacy, AI in Education, Conversational AI, Chatbot, Student-AI Collaboration

1. INTRODUCTION

Artificial intelligence (AI) has rapidly evolved, becoming an integral part of daily life and transforming various sectors, from automating routine tasks to enabling complex decision-making processes. As AI continues to shape industries, communication, and education, there is a growing recognition that understanding and engaging with AI is no longer optional but essential [6]. Just as digital, scientific, and data literacies have become fundamental in the modern world, AI literacy is emerging as a critical competency necessary for navigating an AI-driven society [8, 12].

In response to this need, AI education has gained increasing attention, aiming to equip individuals not only with an understanding of AI concepts but also with the ability to evaluate, apply, and create AI-powered systems. AI literacy extends beyond passive interaction, empowering students to become critical thinkers, problem solvers, and informed decision-makers who can harness AI effectively while considering its broader ethical and societal implications [11, 4]. Ensuring students develop AI literacy is crucial for preparing them to engage with AI responsibly and innovatively in academic, professional, and personal contexts.

Recognizing these benefits, researchers have increasingly explored the impact of AI education on student learning and engagement. Recent research highlights that early AI education not only enhances students' understanding of AI but also promotes interest in AI-related careers [1, 3] and reduces misconceptions about AI's capabilities and limitations [7, 9, 5].

As efforts to develop AI education programs continue to expand, Project-Based Learning (PBL) has emerged as a particularly effective pedagogical approach that allows students to engage with AI in a hands-on, exploratory manner. PBL has been widely recognized for fostering deeper engagement and comprehension, particularly when students develop functional AI applications [10]. Among these applications, chatbot development stands out as an engaging, interactive, and interdisciplinary approach to teaching AI literacy across various subjects. It allows students to apply AI concepts, train machine learning models, and explore human-AI interactions in meaningful ways [13].

A notable example of such application is "AI Made By You" (AMBY), a conversational agent development platform designed to support middle school students in building their own chatbots. Through AMBY, students engage in an it-

erative learning process, refining their chatbot's responses while developing a deeper understanding of AI. Prior studies on AMBY indicate that students gain confidence in their AI-related abilities and develop a stronger interest in AI applications [14]. However, a key challenge lies in the quality of student-created chatbots, which varies significantly depending on students' ability to generate diverse and high-quality training and testing phrases [2]. Furthermore, few studies have explored effective methods for measuring students' AI learning gains across different data sources and mapping these gains to an AI literacy competency framework. Lastly, understanding the relationship between AI literacy and subject-specific knowledge could provide valuable insights for designing AI learning experiences in authentic educational contexts.

To address these challenges, this study proposes enhancing AMBY with Generative AI (GenAI) to scaffold students' chatbot development. Specifically, this research seeks to explore the following questions:

RQ1. How can GenAI enhance students' ability to design high-quality chatbots by improving the diversity, richness, and effectiveness of training phrases?

RQ2. How can students' AI learning gains be measured and mapped to AI literacy competencies using dialogue and log data?

RQ3. In what ways does integrating GenAI into chatbot development enhance students' AI literacy while deepening their understanding of subject-specific disciplinary knowledge?

2. PROGRESS TO DATE
2.1 AMBY: AI Made By You
AMBY is a Conversational chatbot development platform designed for middle school students to explore AI by creating their own conversational agents. As shown in Figure 1, it features a two-panel interface, consisting of a development panel and a testing panel, allowing students to iteratively design, train, and refine their chatbots. Within the development panel, students define chatbot intents, create training phrases and responses (Figure 2). The system supports several types of intents, including top-level intents, which define broad user interactions, as well as follow-up intents, which enable multi-turn conversations. AMBY also includes default and greeting intents to allow students to personalize chatbot responses [13].

To improve chatbot interaction, for each intent, students could design their own training phrases and responses—examples of potential user inputs that train the AI to recognize diverse ways of asking similar questions. They also create testing phrases, which allow them to evaluate chatbot performance and identify areas where responses need refinement. The testing panel provides a simulated chat interface, enabling students to interact with their chatbot, analyze its behavior, and iterate on improvements. This iterative development process mirrors real-world AI training methodologies, helping students develop essential skills in AI problem-solving, debugging, and response optimization. In addition, students

can also customize the agent's voice by selecting between female or male, adjusting the pitch, and modifying the speaking rate to personalize their chatbot's auditory experience.

Previous studies have demonstrated that AMBY enhances students' ability beliefs, AI knowledge, and interests [14]. However, the quality of student-created chatbots varies widely, particularly in terms of productivity (number of training phrases), content richness (word density), and lexical variation (diversity of word choices) [2]. Some students excel in developing well-structured, contextually rich chatbot interactions, while others struggle to generate meaningful training data, resulting in chatbots with limited conversational depth. This variability underscores the need for real-time scaffolding and feedback mechanisms to support students in improving chatbot quality.
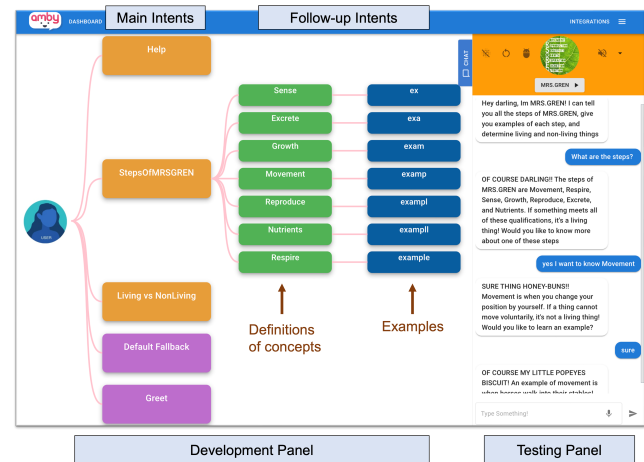

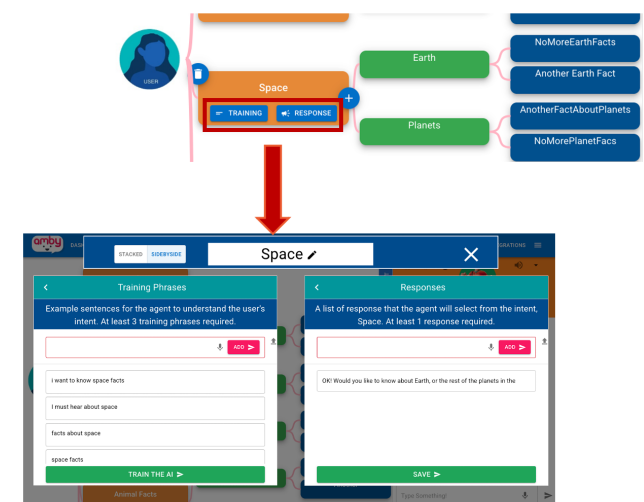**Figure 1: The Interface of AMBY, with an Example Chatbot**

Figure 2: Setting up Training Phrases and Responses

2.2 Generative AI Integration in AMBY
To address the challenges in chatbot quality and foster AI literacy development, this study proposes integrating GenAI

into AMBY's chatbot development workflow. By embedding real-time GenAI-powered scaffolding, AMBY can provide students with dynamic feedback, content enrichment suggestions, and lexical variation prompts as they develop training phrases, responses, and testing for their chatbots. An important note is that we will implement a safeguard filter into the generative AI system to prevent it from providing any responses that are inappropriate for K–12 students.

Specifically, integrating GenAI-driven support into AMBY can enhance chatbot development in several key ways. First, AI-assisted phrase generation can provide students with more examples of training phrases to help them construct more comprehensive chatbot responses. By analyzing students' inputs and providing appropriate suggestions, GenAI can assist students in creating a more diverse set of training phrases, improving chatbot recognition of varied conversational patterns. Second, content enrichment guidance can encourage students to refine chatbot responses by adding specific details, structured explanations, and contextual depth. This scaffolding helps students develop more meaningful and informative chatbot interactions, aligning with research on adaptive learning environments that personalize instruction based on student needs. Third, lexical diversity recommendations can prompt students to explore synonyms, alternative phrasing, and varied sentence structures, ensuring that their chatbots can recognize and respond to a broader range of student inputs. Fourth, GenAI can provide targeted debugging support when students encounter errors during chatbot testing. Specifically, GenAI can help students identify which intent or follow-up intent may be incorrectly configured and offer guidance that facilitates troubleshooting and supports deeper understanding of chatbot logic.

Moreover, GenAI can act as a collaborative peer by engaging students in an interactive co-design process. Students can iteratively brainstorm, test, and refine training phrases and responses with GenAI as a responsive design partner, helping them expand their chatbot's conversational scope. This also reduces the time required to teach students how to use AMBY, allowing them to access and utilize the platform anytime and anywhere at their convenience. Lastly, this research also aims to explore how AI-powered feedback can support students in developing AI concepts across disciplinary subjects. For instance, in science education, students could train scientific chatbots to simulate scientific explanations, such as describing biological processes like photosynthesis, classifying organisms, or modeling climate change effects. In mathematics, chatbots could help students solve algebraic equations, generate step-by-step explanations for problem-solving, or provide interactive word problem practice by adapting responses based on students' inputs. In computer science, chatbots could assist in debugging code, explaining fundamental programming concepts like loops and conditionals, or even guiding students through algorithmic problem-solving. By integrating GenAI into AMBY, students can apply AI concepts in meaningful ways across different disciplines, deepening their understanding of both AI literacy and subject-specific knowledge.

3. OPPORTUNITIES AND CHALLENGES
The integration of Generative AI into AMBY brings both valuable opportunities and notable challenges. One key op-

portunity is leveraging self-regulated learning (SRL) principles to refine GenAI-driven scaffolding so that students receive structured, goal-oriented support while minimizing AI-generated hallucinations. SRL, which encompasses the phases of planning, monitoring, and reflecting, is known to help students set goals, track their progress, and evaluate their learning strategies [15], making it a strong theoretical foundation for designing AI support that promotes autonomy and metacognitive awareness. By aligning feedback with students' learning processes, GenAI can guide students through iterative chatbot development, encouraging reflection, self-monitoring, and strategic problem-solving. However, ensuring that AI-generated feedback remains accurate, relevant, and pedagogically sound remains a key challenge. Another challenge is effectively measuring AI literacy using dialogue interactions and log data. Furthermore, mapping students' AI learning gains to established AI literacy competency frameworks such as Long and Magerko's framework [8], remains an underexplored area, particularly in K-12 education.

Lastly, this study aims to examine the relationship between AI literacy and subject-domain learning gains as students develop subject-specific chatbots. While chatbot development fosters a project-based, interdisciplinary AI learning experience, its impact on disciplinary knowledge in science, mathematics, literature, and computer science is still unclear. Investigating how AI literacy intersects with subject-specific learning outcomes will provide insights into the broader educational implications of AI integration.

4. CONCLUSION
As AI literacy becomes an increasingly critical skill, project-based chatbot development provides a hands-on and interactive approach for students to engage with AI concepts. However, inconsistencies in the quality of student-created chatbots highlight the need for effective scaffolding to support students in designing more robust conversational agents. This study explores the role of GenAI-driven scaffolding in enhancing students' productivity, content richness, and lexical diversity during chatbot development. By incorporating real-time feedback and interdisciplinary AI applications, this approach aims to strengthen AI literacy instruction while enhancing domain-specific learning. Future work will focus on refining GenAI's scaffolding mechanisms, aligning chatbot development components with AI literacy competencies, and examining the connection between AI literacy and subject-specific learning gains in K-12 education.

5. ACKNOWLEDGMENTS

6. REFERENCES
[1] A. Bewersdorff, X. Zhai, J. Roberts, and C. Nerdel. Myths, mis-and preconceptions of artificial intelligence: A review of the literature. *Computers and Education: Artificial Intelligence*, 4:100143, 2023.

[2] C. Borchers, X. Tian, K. E. Boyer, and M. Israel. Combining log data and collaborative dialogue features to predict project quality in middle school ai education. In *Under review*, 2025.

[3] S. Cave, K. Coughlan, and K. Dihal. " scary robots" examining public responses to ai. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, pages 331–337, 2019.

[4] T. K. Chiu, Z. Ahmad, M. Ismailov, and I. T. Sanusi. What are artificial intelligence literacy and competency? a comprehensive framework to support them. *Computers and Education Open*, 6:100171, 2024.

[5] M. Ghallab. Responsible ai: requirements and challenges. *AI Perspectives*, 1(1):1–7, 2019.

[6] M. Kandlhofer, G. Steinbauer, S. Hirschmugl-Gaisch, and P. Huber. Artificial intelligence and computer science in education: From kindergarten to university. In *2016 IEEE frontiers in education conference (FIE)*, pages 1–9. IEEE, 2016.

[7] K. Kim, K. Kwon, A. Ottenbreit-Leftwich, H. Bae, and K. Glazewski. Exploring middle school students' common naive conceptions of artificial intelligence concepts, and the evolution of these ideas. *Education and Information Technologies*, 28(8):9827–9854, 2023.

[8] D. Long and B. Magerko. What is ai literacy? competencies and design considerations. In *Proceedings of the 2020 CHI conference on human factors in computing systems*, pages 1–16, 2020.

[9] P. Mertala and J. Fagerlund. Finnish 5th and 6th graders' misconceptions about artificial intelligence. *International Journal of Child-Computer Interaction*, 39:100630, 2024.

[10] D. T. K. Ng, M. Lee, R. J. Y. Tan, X. Hu, J. S. Downie, and S. K. W. Chu. A review of ai teaching and learning from 2000 to 2020. *Education and Information Technologies*, 28(7):8445–8501, 2023.

[11] D. T. K. Ng, J. K. L. Leung, S. K. W. Chu, and M. S. Qiao. Conceptualizing ai literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2:100041, 2021.

[12] M. Pinski and A. Benlian. Ai literacy for users–a comprehensive review and future research directions of learning methods, components, and effects. *Computers in Human Behavior: Artificial Humans*, page 100062, 2024.

[13] X. Tian, A. Kumar, C. E. Solomon, K. D. Calder, G. A. Katuka, Y. Song, M. Celepkolu, L. Pezzullo, J. Barrett, K. E. Boyer, et al. Amby: A development environment for youth to create conversational agents. *International Journal of Child-Computer Interaction*, 38:100618, 2023.

[14] X. Tian, Y. Song, S. Zhang, P. G. Prasad, W. S. Ménard, C. E. Solomon, C. F. Wise, E. Dobar, T. McKlin, K. E. Boyer, and M. Israel. "my bot talks science!": Integrating conversational ai into middle school science classrooms through chatbot development. *Under review*, 2025.

[15] B. J. Zimmerman. Becoming a self-regulated learner: An overview. *Theory into practice*, 41(2):64–70, 2002.