Majumdar, R. et al. (Eds.) (2025). Proceedings of the 1st International Conference on Learning Evidence and Analytics. Asia-Pacific Society for Computers in Education

Effectiveness of Instruction on the Utilization of Generative AI in Problem Solving: A Case Study of Japanese Junior High School

Koki TAMAKI[a*], Shota HIGUCHI[b], Akira MAEDA[c] & Masaki RAKU[d]

[a]*Research Center for Higher Education, Tokushima University, Japan*
[b]*Department of Mathematics and Science Education, Tokyo University of Science, Japan*
[c]Graduate School of Systems Information Science*, Future University Hakodate, Japan*
[d]*Psychology, Meiji Gakuin University, Japan*
* tamaki.kouki@tokushima-u.ac.jp

Abstract: This study investigated the effectiveness of instructional practices involving generative AI (ChatGPT) in supporting problem-solving activities among junior high school students. The instructional design included basic usage of generative AI and its application in real-world problem-solving contexts. The results of measuring changes in students' ChatGPT literacy using the ChatGPT Literacy Scale showed improvements in all subcategories. In addition, analysis of prompt quality and usage patterns revealed that students became more sophisticated in their interactions with the AI. These findings suggest that problem-solving activities using generative AI can contribute to the development of students' AI literacy. However, the study also identified several challenges, including limited instructional time, individual differences in AI literacy, and the need for more diverse evaluation methods. Moving forward, the study highlights the importance of continuous integration of AI into the curriculum, the use of learning analytics for deeper assessment, and the design of adaptive instructional support tailored to diverse learners.

Keywords: AI Literacy, Generative AI, Problem-solving, AI Education

1. Introduction

The rapid advancement of generative AI has led to its widespread adoption across various domains. In response to this trend, the OECD (2023) emphasizes the need for new forms of instruction that enable students to use generative AI effectively. Similarly, Japan's Ministry of Education, Culture, Sports, Science and Technology (MEXT, 2023) has published provisional guidelines on the use of generative AI in elementary and junior high schools, offering direction for its appropriate integration into education.

Recently, in order to provide appropriate education on AI literacy perspectives to learners, a number of studies have sought to establish assessment frameworks for AI literacy. For example, Lee and Park (2024) developed the "ChatGPT Literacy Scale" to promote the healthy and effective use of generative AI. This scale evaluates the competencies required for using ChatGPT across five distinct dimensions. Laupichler et al. (2023) also created the "Scale for the Assessment of Non-Experts' AI Literacy," a tool designed to measure personal AI literacy among non-experts, composed of three factors.

In this way, efforts are underway to evaluate the AI literacy of individuals. This is driven by the widespread use of generative AI in everyday life. In fact, children often turn to generative AI as a solution when faced with challenges in their daily lives. A survey by Forman et al. (2023) of high school students revealed that generative AI is frequently used in non-academic contexts such as information search and entertainment.

Despite students' frequent use of generative AI outside of school, it remains unclear how appropriately they are using it. Furthermore, the effectiveness of school-based instruction on generative AI and its impact on students' ability to use these tools is not yet well understood.

To address this gap, the present study investigates the effects of instructional interventions aimed at promoting generative AI use in problem-solving contexts among junior high school students. By providing both general guidance on using generative AI and specific strategies for applying it to problem-solving tasks, this study aims to explore how such instruction may influence students' literacy and usage patterns. This is a crucial step toward the effective integration of generative AI into future educational practices.

2. Purpose

The purpose of this study is to examine whether junior high school students' ChatGPT literacy and usage improve as a result of instruction on generative AI and its application to problem-solving activities. Specifically, the study addresses the following research questions:
(1) Does students' ChatGPT literacy improve after instruction?
(2) How does the content of students' prompts to ChatGPT change?
(3) How does the way students evaluate and utilize ChatGPT's output change?

3. Methodology

3.1 Participants

This study was conducted with 41 second-year junior high school students from Japan, who participated in a special lecture. Before the implementation, an explanation of the research outline and data handling procedures was provided, and only students who consented to the data collection were included in the study. Additionally, consent was obtained from the students' guardians for the educational use of ChatGPT.

3.2 Procedure

The implementation took place in February 2025. The lessons were structured into 50-minute sessions, with two sessions in the first week and one session in the second week, totaling three sessions (150 minutes). The flow of the practice is shown in Figure 1.

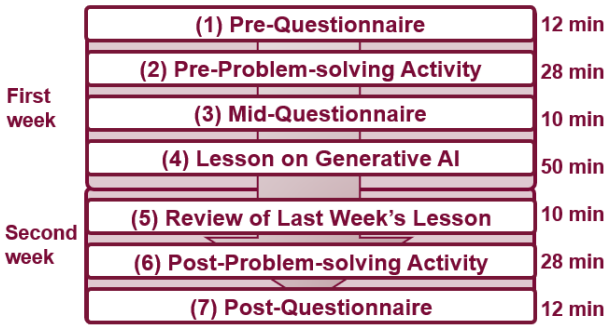

Figure 1. Procedure.

First, (1) a pre-questionnaire was conducted to assess the students' ChatGPT literacy. Then, (2) the problem-solving activity was conducted. This activity lasted 25 minutes, followed by a 3-minute period for summarizing the results. The task was to propose something that would enhance school life, with a budget limit of 100,000 yen and a requirement that it could be set up within two months. The activity was carried out in groups of four students, with one iPad per group, allowing them to use generative AI for assistance as needed.

After that, (3) a mid-questionnaire was conducted to assess the students' self-evaluation of their problem-solving. Then, (4) a basic lesson on generative AI was provided. The lesson covered how to input prompts to generative AI, the risks of hallucinations and misinformation, and how to use generative AI in the problem-solving process. Special emphasis was placed on using generative AI not just as an alternative for information search but also as a problem-solving advisor, helping students review their proposed solutions.

In the second week, we conducted (5) a review of last week's lesson. Then,(6) a post-problem-solving activity was carried out. The conditions were the same as in the pre-activity, but the task was changed. The new task asked students to propose a school event that would help them discover new aspects of their friends. The conditions were that the event should be completed within two days and with a budget of 100,000 yen. Finally, (7) a post-questionaire was conducted to assess the students' ChatGPT literacy and confidence in the problem-solving activity.

3.3 Chat tools with generative AI

In this study, a system was developed for students to interact with generative AI while solving problems. The system consists of (1) a dialogue function, (2) dialogue management function, and (3) data export function (Figure 2). Students enter questions or opinions into the system via a text input form, and the system uses OpenAI's GPT-4o-mini model via an API. The system does not involve any special prompt tuning for the model, meaning the role of problem-solving facilitator or coach is not explicitly instructed. The input from the students is sent to ChatGPT-4o-mini via the API, and the model generates responses considering the context of the ongoing conversation. To ensure that the generative AI considers the context of the dialogue, the past conversation history is included in the prompt.

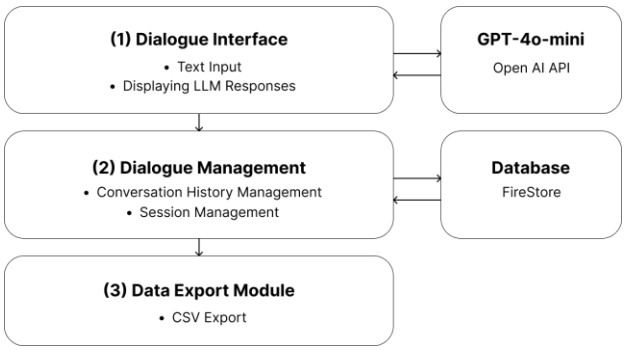

Figure 2. System Design.

3.4 Data Collected

3.4.1 Input to AI

The system records the interactions between the students and generative AI, saving the following information:

- Student input
- Response from GPT-4o-mini
- Date and time of the student's input and AI's response
- Session ID (to identify consecutive dialogues from the same group of students) These data are stored in FireStore and can be exported in CSV format for analysis.

3.4.2 Questionnaire

In this study, we conducted the "ChatGPT Literacy Scale," which consists of 25 items across five domains considered necessary for effective use of generative AI. For this study, we selected only the three domains most relevant to our research objectives: Critical Evaluation (CE), Communication Proficiency (CP), and Creative Application. These were assessed using a 5-point Likert scale and administered both before and after the intervention.

3.4.3 Dialogue Recordings

The students' dialogues during the problem-solving activity were recorded. One iPad was placed at the center of each group to observe and analyze how the students engaged in the problem-solving activity.

4. Results and Discussion

4.1 Questionnaire

We performed paired t-tests and calculated effect sizes using Cohen's d. The results of the ChatGPT Literacy Scale are shown in Table 1 (*M* = mean, *SD* = standard deviation). Improvements were observed across all components: Critical Evaluation (CE), Communication Proficiency (CP), and Creative Application (CA). These findings suggest that both foundational instruction on generative AI and guidance on how to utilize it during problem-solving activities can enhance ChatGPT literacy.

Table 1. *Pre- and post-comparison of ChatGPT Literacy Scale*

Factor	Pre		Post		M_{post}-M_{pre}	*t*		*d*
	M	*SD*	*M*	*SD*				
Critical Evaluation	3.38	0.57	3.69	0.72	0.31	2.46	*	0.40
Communication Proficiency	3.18	0.73	3.72	0.81	0.54	3.69	***	0.60
Creative Application	3.37	0.79	3.77	0.78	0.40	2.65	*	0.43
n=38, 5 Likert scale								**p < 0.05, ***p < 0.001*

4.2 Evaluation of input to AI

In the pre-problem-solving activity, many groups used generative AI primarily to generate ideas. This aligns with previous findings that generative AI is often used as a tool for information retrieval (Forman et al., 2023). However, some groups (e.g., Group 1) attempted to redefine the problem itself through their prompts. These differences suggest a wide variation in students' literacy regarding generative AI.

In the post-problem-solving activity, new patterns of generative AI use emerged that were not observed before. For instance, initially verbose prompts became more concise (Group 1), and learners refined their prompts to better the output to their needs (Group 6).

Additionally, a group that previously struggled to formulate questions (Group 9) began using generative AI when their own ideas were lacking. Another group (Group 8) repeatedly generated outputs to refine ideas that aligned closely with their goals. These behaviors were not present in the pre-task phase.

Overall, these findings suggest that students were able to use generative AI more effectively for problem-solving after receiving instruction.

4.3 Evaluation of output to AI

In the pre-problem-solving activity, most groups used the output generated by the generative AI without critically evaluating its content. However, after the instruction, several groups (Groups 2, 6, and 9) were observed to approach the AI-generated responses with a more skeptical attitude, engaging in discussions about the validity and appropriateness of the output.

In particular, Group 9 not only referenced the AI's responses but also discussed missing perspectives and alternative ideas, using the output as a springboard for collaborative

problem-solving. These findings suggest a shift from passive consumption of AI-generated content to more active and critical engagement with the technology, fostering meaningful dialogue and cooperative problem-solving among students.

5. Conclusion

This study examined the effects of instructional practices that integrated generative AI into problem-solving activities among junior high school students, focusing on changes in their ChatGPT literacy and AI usage. The findings indicate that even relatively short-term interventions can enhance learners' awareness and strategic use of generative AI in problem-solving. In particular, learners began to adopt more intentional prompting strategies, and some learners showed improvement in evaluating AI-generated responses and integrating them into group discussions. These results emphasize the potential of educational curricula to foster the literacy necessary for effective AI use.

At the same time, several challenges emerged. First, there were limitations in both instructional time and content. Because of the restricted schedule, the instruction had to cover both the basic usage of generative AI and its application to problem-solving, which limited opportunities for iterative practice and reflection. While the positive short-term outcomes are encouraging, it remains unclear whether these improvements can be sustained over time or transferred to other contexts.

In addition, the diversification of assessment methods is a key issue. Although this study primarily relied on questionnaire surveys and activity records, deeper insights could be gained by incorporating a mixed-methods evaluation, including interviews, qualitative analysis of prompts, and learning log analysis. For example, analyzing learning logs such as the length and syntactic complexity of prompts, changes in keywords and specificity, and the frequency or patterns of prompt revision for the same task, would allow for behavioral-level tracking of literacy development.

Moreover, by analyzing which stages of the problem-solving process (e.g., information gathering, idea generation, evaluation) students utilize AI, one can align their behavior with existing models and model the relationship between AI use and problem-solving quality to identify typical patterns. Furthermore, clustering learners based on the quality and frequency of prompts or their level of reliance on AI could inform the design of adaptive instructional strategies, offering an important step toward personalized AI-enhanced education.

As for future directions, expanding practice to other age groups and learning contexts is promising. While this study focused on junior high school students, similar approaches could extend to elementary or high school learners, or to other subjects and inquiry-based activities. As generative AI evolves, fostering AI literacy—not just the ability to use AI, but the critical competencies to interpret, assess, and apply its outputs across contexts—becomes increasingly important. Learning analytics is regarded as essential in this context.

References

Forman, N., Udvaros, J., & Avornicului, M. S. (2023). ChatGPT: A new study tool shaping the future for high school students. *Future, 5*(6), 7. https://doi.org/10.59287/ijanser.2023.7.4.562

Laupichler, M. C., Aster, A., Haverkamp, N., & Raupach, T. (2023). Development of the "Scale for the assessment of non-experts' AI literacy"–An exploratory factor analysis. *Computers in Human Behavior Reports*, *12*, 100338. https://doi.org/10.1016/j.chbr.2023.100338

Lee, S., & Park, G. (2024). Development and validation of ChatGPT literacy scale. *Current Psychology*, *43*(21), 18992-19004. https://doi.org/10.1007/s12144-024-05723-0

MEXT (2023). Temporary Guidelines for Use of Generative AI in Primary and Secondary Education. https://www.mext.go.jp/content/20241226-mxt_shuukyo02-000030823_001.pdf

OECD (2023). OECD Digital Education Outlook 2023 Towards an Effective Digital Education Ecosystem. https://www.oecd.org/en/publications/oecd-digital-education-outlook-2023_c74f03de-en/full-report/emerging-governance-of-generative-ai-in-education_3cbd6269.html