

***Combining Realistic Story Situations and GPT-Based NPC Framework for
Historical Knowledge Problem-Solving Games***

Cheng-Yuan Wei, National Taiwan University of Science and Technology, Taiwan
Yu-Chi Chen, National Taiwan University of Science and Technology, Taiwan
Chih-Chung Chien, National Taiwan University of Science and Technology, Taiwan
Huei-Tse Hou, National Taiwan University of Science and Technology, Taiwan

The Asian Conference on Education 2024
Official Conference Proceedings

Abstract

In recent years, the application of generative AI's (GAI) adaptive characteristics in education has rapidly emerged. However, general GAI in educational games often lacks realistic situations and operational fidelity, leading to limited authentic experiences and difficulties in learning transfer. Additionally, GAI frequently produces inaccurate or off-topic responses. To address these limitations, this study designs a historical education problem-solving game based on a previous research team's framework of using GPT as a game Non-Player Character (NPC) (Chen & Hou, 2024). The game incorporates realistic story situations and reduces off-topic NPC responses. In the game, learners play as the close friend of the protagonist and collaborate with GPT-based NPC peers designed by the research team. They conduct online information searches within a limited time to find the historical period in which the game protagonist disappeared and the name of the current building to save the protagonist. A total of 17 participants engaged in the empirical evaluation of this study. The study found that learners exhibited a high flow state, perceived the game as highly playable and enjoyable, and had a high acceptance of the game. Additionally, 70% of learners believed that the game helped with historical learning, and nearly 50% felt that it was closer to real human interaction compared to typical GPT conversations. The study demonstrates that a GPT-based NPC, enhanced with contextual stories and reduced off-topic responses, can effectively improve learners' gaming experience.

Keywords: Generative AI, Problem-Solving Exploration Skills, Operational Realism

iafor

The International Academic Forum
www.iafor.org

Introduction

With the rapid emergence of generative AI, various educational applications have quickly risen, bringing benefits to learners. Research has shown that AI-generated gamified assessment strategies significantly enhance student engagement and motivation, improving learning outcomes (Bachiri et al., 2023). Integrating escape room games with AI chatbots allows educators to design immersive educational experiences more easily, enhancing teaching efficiency and effectiveness (Fotaris et al., 2023). AI-assisted gamified learning can boost students' intrinsic motivation, reduce cognitive load, and foster effective learning behaviors in scientific education (Chen, C.-H., & Chang, C.-L., 2024). Chen and Hou (2024) found that incorporating ChatGPT-based situational NPCs into educational games enables students to maintain high motivation and engagement without excessive anxiety. However, general generative AI lacks situational and operational realism in educational games, limiting the sense of authenticity and making learning transfer more difficult. For instance, Benjamin Emihovich and colleagues (2020) examined improvements in problem-solving skills and found no significant difference in PISA problem-solving test performance between using popular video games like *World of Warcraft* and cognitive training games like *CogniFit*. This suggests that without realistic scenarios or relevant subject knowledge, effectiveness remains limited, and the same holds for AI-driven research. Furthermore, current generative AI answers often contain a certain error rate, but tutor versions designed by researchers show lower error rates than those without prior configuration (Bastani et al., 2024).

Building on this, our research developed a narrative context and implemented character-personalized GPT. Furthermore, based on previous research by Chen and Hou (2024), which used GPT as an NPC in games, we designed a historical educational problem-solving game titled *The Disappearance of Bi-Ting*, incorporating a realistic narrative and reducing off-topic NPC responses.

Methods

Participants in this study were 17 adults over the age of 18 from Taiwan (8 males and 9 females), recruited through online registration. None of the participants had previously experienced a GenAI-integrated game-based learning activity. Each participant used a personal computer and conducted the online learning activity in a private space. The entire study was conducted online.

This activity is conducted on a Google Sites webpage, integrating game text, Google Form puzzles, and GPT-based NPCs. Before starting, learners watch an introduction video covering the game story and instructions. Learners play as the protagonist's close friend, tasked with finding the year and specific time-space location where their friend has disappeared within a limited timeframe. Throughout the game, learners interact with a GPT-based NPC developed by the team, which guides players to think critically without directly providing correct answers. Learners must assess the NPC's information and may conduct online searches for further information.

The activity was divided into several stages. First, a 10-minute introduction was given on the game story and operation. This was followed by a 30-minute game experience. After the game, a 20-minute post-test was conducted. Finally, a 10-minute discussion and explanation session was held to conclude the day's activity, making the total duration approximately 70 minutes.

The instruments included a Flow Scale, a Game Anxiety Scale, a Game Feedback Scale, a Game Elements Scale, and a self-reported questionnaire. The Flow Scale, translated and revised by Hou and Li (2014) from Kill's Flow Scale (2006), covered flow antecedents and flow experience. To assess participant anxiety, this study adapted Hung's (2001) Activity Anxiety Scale. Game Feedback was measured using Davis's (1989) scale, which includes perceived usefulness and ease of use. The Game Elements Scale was based on game motivational elements proposed by Hou (2016). All of the above scales were evaluated using a five-point Likert scale. The self-reported questionnaire collected data on realism and usage effectiveness based on the characteristics of the GPT-based NPC and historical learning outcomes in this study.



Figure 1: Game Format and NPC Interaction Interface

Results and Discussions

In the quantitative questionnaire section of this study, a one-sample Wilcoxon signed-rank test was used to analyze. Table 1 shows that the overall flow ($M=4.19$) is significantly higher than the median of the scale (i.e., 3), indicating that the design of this study provided learners with clear goals, appropriate difficulty, and a high sense of self-mastery, thereby fostering a high flow experience and encouraging active participation. Game anxiety ($M=2.80$) did not significantly differ from the median of the scale (i.e., 3), suggesting that the study design created a moderate level of anxiety during the learning process. Game acceptance ($M=4.08$) was significantly higher than the median of the scale (i.e., 3), indicating that the game is easy to engage with and helps users develop online information skills. Additionally, game elements ($M=4.19$) were significantly higher than the median of the scale (i.e., 3), suggesting that the study design allowed learners to feel a sense of control while incorporating a degree of uncertainty and challenge, which made the game enjoyable. Participants also reported a sense of achievement upon completing levels and expressed a willingness to experience similar games in the future.

Table 1: Flow and Activity Anxiety and Game Acceptance Descriptive Statistical Analysis

| Dimension | <i>M</i> | <i>SD</i> | <i>Z</i> | Sig. |
|----------------------|----------|-----------|----------|-------|
| Overall Flow | 4.19 | 0.62 | 3.62*** | 0.000 |
| Flow antecedents | 4.18 | 0.65 | 3.52*** | 0.000 |
| Flow experience | 4.20 | 0.64 | 3.58*** | 0.000 |
| Game Anxiety | 2.80 | 0.82 | -1.11 | 0.266 |
| Game Feedback | 4.08 | 0.49 | 3.52*** | 0.000 |
| Game Usefulness | 3.93 | 0.64 | 3.45** | 0.001 |
| Game Ease of Use | 4.27 | 0.63 | 3.54*** | 0.000 |
| Game Elements | 4.19 | 0.65 | 3.64*** | 0.000 |

** $p < 0.01$, *** $p < 0.001$

In the self-reported questionnaire section, as shown in Figure 2 and 3, regarding the realism of NPC dialogue, 41.8% of participants felt that it resembled a real person, while 11.76% mentioned that it was similar to other AIs but still felt like conversing with a human. They noted that the responses were more emotional, less rigid, and did not reveal all the answers at once, prompting further inquiries. In terms of providing informational support, 76.47% reported that the NPC was helpful in the information search process, such as by answering with relevant details (who, when, where, what), offering related keywords, and helping to narrow down the scope of clues. During gameplay, 70.59% indicated that with the help of NPC Xiao-Rou, they could conduct information searches more quickly and efficiently focus on relevant data. Additionally, 70.59% of participants mentioned that the game enhanced their understanding of Tainan's history during the Japanese colonial period, motivating them to explore the bustling locations, historical events, architecture, and streets of Tainan at that time, thus contributing positively to their learning.

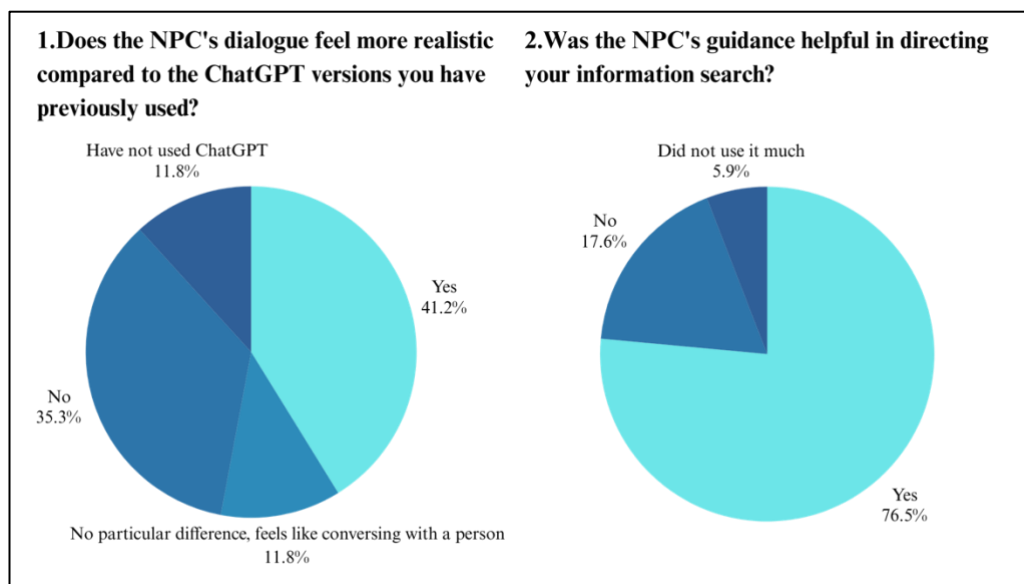


Figure 2: Responses of Participants in the Self-Reported Questionnaire

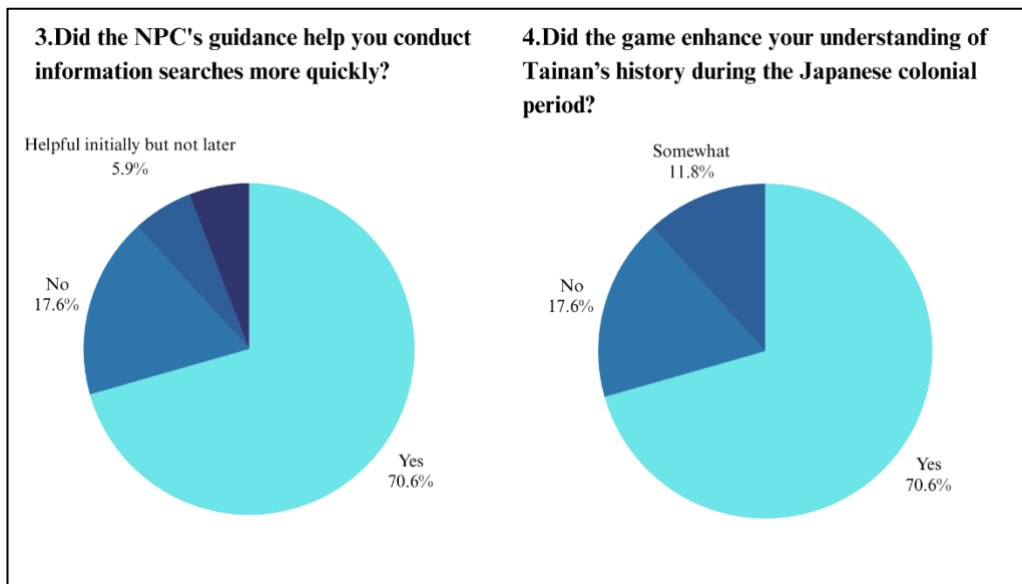


Figure 3: Responses of Participants in the Self-Reported Questionnaire

Conclusions and Limitations

“The Disappearance of Bi-Ting” is a history-themed problem-solving game, based on situational storytelling and GPT-based NPCs designed to minimize off-topic responses. Players use textual clues and interact with the NPC to gradually gather and focus on relevant information, fostering participants' information-gathering skills. Data analysis reveals significant differences in flow state and game acceptance, suggesting that the problem-solving game, which integrates realistic storytelling and GPT-based NPCs, not only enhances learners' gaming experience and flow but also improves their ability to utilize information effectively. Additionally, anxiety levels did not reach statistical significance, indicating that the study did not induce excessive or inadequate activity-related anxiety among participants. This approach can serve as a model in designing educational games. The study's historical knowledge problem-solving game was reported by learners to be beneficial for understanding local historical culture.

Future research will employ a quasi-experimental design to compare the effects on learners between an experimental group using GPT-based NPCs and a control group (two participants per group engaging in online text discussions without GPT-based NPCs).

Acknowledgments

This research was supported by the “Empower Vocational Education Research Center” of National Taiwan University of Science and Technology (NTUST) from the Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan and projects from the Ministry of Science and Technology, Taiwan, under contract number MOST- 110-2511-H-011 -004 -MY3 and MOST-111-2410-H-011 -004 -MY3.

References

- Bachiri, Y.-A., Mouncif, H., & Bouikhalene, B. (2023). Artificial Intelligence Empowers Gamification: Optimizing Student Engagement and Learning Outcomes in E-learning and MOOCs. *International Journal of Engineering Pedagogy (IJEP)*, 13(8), 4–19. <https://doi.org/10.3991/ijep.v13i8.40853>
- Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakçı, O., & Mariman, R. (2024). Generative ai can harm learning. *Available at SSRN*, 4895486. <https://dx.doi.org/10.2139/ssrn.4895486>
- Chen, C.-H., & Chang, C.-L. (2024). Effectiveness of AI-assisted game-based learning on science learning outcomes, intrinsic motivation, cognitive load, and learning behavior. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-024-12553-x>
- Chen, Y.-C., & Hou, H.-T. (2024). A Mobile Contextualized Educational Game Framework With ChatGPT Interactive Scaffolding for Employee Ethics Training. *Journal of Educational Computing Research*. <https://doi.org/10.1177/07356331241268505>
- Chen, Y.-C., & Hou, H.-T. (2024). *Design and Evaluation of the Framework of ChatGPT-Based Situational Non-Player Character Scaffolding for Educational Games*. 720–721. <https://doi.org/10.1109/iiia-aa63651.2024.00152>
- Chou, Y.-S., Hou, H.-T., Chang, K.-E., & Su, C.-L. (2021). Designing cognitive-based game mechanisms for mobile educational games to promote cognitive thinking: an analysis of flow state and game-based learning behavioral patterns. *Interactive Learning Environments*, 1–18. <https://doi.org/10.1080/10494820.2021.1926287>
- Emihovich, B., Roque, N., & Mason, J. (2020). Can Video Gameplay Improve Undergraduates' Problem-Solving Skills? *International Journal of Game-Based Learning*, 10(2), 21–38. <https://doi.org/10.4018/ijgbl.2020040102>
- Fotaris, P., Mastoras, T., & Lameris, P. (2023). Designing Educational Escape Rooms with Generative AI: A Framework and ChatGPT Prompt Engineering Guide. *European Conference on Games Based Learning*, 17(1), 180–189. <https://doi.org/10.34190/ecgbl.17.1.1870>

Contact email: hthou@mail.ntust.edu.tw