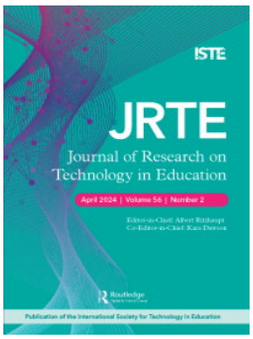

Routledge
Taylor & Francis Group

Journal of Research on Technology in Education

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/ujrt20

Learning by playing with generative AI: design and evaluation of a role-playing educational game with generative AI as scaffolding for instant feedback interaction

Chih-Chung Chien, Hung-Yu Chan & Huei-Tse Hou

To cite this article: Chih-Chung Chien, Hung-Yu Chan & Huei-Tse Hou (08 Apr 2024): Learning by playing with generative AI: design and evaluation of a role-playing educational game with generative AI as scaffolding for instant feedback interaction, Journal of Research on Technology in Education, DOI: 10.1080/15391523.2024.2338085

To link to this article: https://doi.org/10.1080/15391523.2024.2338085

Published online: 08 Apr 2024.

JOURNAL OF RESEARCH ON TECHNOLOGY IN EDUCATION
https://doi.org/10.1080/15391523.2024.2338085

Learning by playing with generative AI: design and evaluation of a role-playing educational game with generative AI as scaffolding for instant feedback interaction

Chih-Chung Chien, Hung-Yu Chan and Huei-Tse Hou

Graduate Institute of Applied Science and Technology, National Taiwan University of Science and Technology, Taipei, Taiwan

ABSTRACT

In this study, an online contextualized educational game was designed to provide interactive simulated dialogues using generative AI scaffolding (using *ChatPDF*) in a contextualized game as scaffolding for immediate feedback, where learners can access guides and explore knowledge. This study analyzed learners' behaviors while performing AI prompting in the interactive scaffolding, as well as learners' psychological responses. A total of 59 students participated in this study. The results of the study showed that learners had significantly high flow and low activity anxiety in the game tasks, while game feedback and scaffolding usefulness had significant effects on learning aids. The generative AI instant feedback interactive scaffolding had a certain high percentage of direct answers or indirect suggestions, which is suitable for interactive scaffolding.

ARTICLE HISTORY
Received 15 September 2023
Revised 18 January 2024
Accepted 4 March 2024

KEYWORDS
Generative AI; ChatGPT; prompting; interactive scaffolding; game-based learning

Introduction

In recent years, the advancement of AI has seen more mature technological developments, allowing generative AI to break through the bottleneck of traditional AI. Generative AI is based on a deep learning model that generates creative and randomized feedback, making replies more flexible. Many studies have begun to explore its impact on teaching and learning, including its natural language advantages (Rahman & Watanobe, 2023). Learners interacting with instant answers in generative AI systems have also been positively evaluated (Sánchez-Ruiz et al., 2023). Because of its simulated dialog features, using generative AI as scaffolding has the potential to provide instant guidance to learners. Especially in the field of game-based learning research to promote learning motivation, it is expected to provide more realistic scaffolding guidance if generative AI is utilized as a non-player character in role-playing educational games. This issue deserves in-depth empirical analysis and investigation.

Generative AI and instructional usage

The use of generative AI in various domains is becoming increasingly emphasized, and this technology can generate a variety of simulated interactions and feedback. Due to its ability to generate answers to questions posed by the user, there are several scholars who worry that AI assistance is leading to cheating concerns (Lukpat, 2023; Stokel-Walker, 2022). Teachers and educational institutions must become more accurately familiar with the use of *ChatGPT* to ensure

CONTACT Huei-Tse Hou hthou@mail.ntust.edu.tw Graduate Institute of Applied Science and Technology, National Taiwan University of Science and Technology, #43 Keelung Road, Section 4, Taipei, Taiwan.

that its learning advantages do not hinder the learning process (Sánchez-Ruiz et al., 2023). Rahman and Watanobe (2023) mentioned that one of the biggest advantages of *ChatGPT* for educational applications is its ability to understand and respond to natural language queries. This makes the *ChatGPT* application more intuitive and easier to use for learning by making it seem as if the learner has a teacher on hand to ask questions. Most of the current studies have examined the strengths and threats of *ChatGPT*, but fewer have explored the strengths and limitations of the learner's question-and-answer interactions when utilizing *ChatGPT* as a learning interactive tool. Sánchez-Ruiz et al. (2023) found that learners in an engineering education study were prone to use *ChatGPT* to achieve generative AI interactions, to simulate real human-like communication, and to experience and learn structured conversational patterns. Meanwhile, learners showed positive confidence in ChatGPT's instant response interactions and learning process usage, rating it highly (Sánchez-Ruiz et al., 2023). It also enhanced the interactive simulation of the game process. However, ChatGPT's generated content currently still has its limitations in terms of accuracy and correctness, and may even produce incorrectly generated content that affects learning.

In 2023, many cloud tools employed the *ChatGPT* API to focus on specific data sources for feedback generation. For example, users can upload a PDF file to a database and chat directly with the generative AI (e.g., *ChatPDF, PDFGPT.IO*), as if an expert were able to answer all of the user's questions about the uploaded file. This mechanism has the potential to provide instant scaffolding in educational games because it can focus more on the learning content in the PDF to generate feedback. This is expected to improve the accuracy and correctness of the generated information.

Game-based learning

Game-based learning (GBL) has been defined as a learner-centered approach to learning while integrating the entertainment of games with learning (Prensky, 2003). Many studies have indicated that game-based learning might improve learning motivation and learning effectiveness (Bawa et al., 2018; Chang et al., 2017). Chou et al. (2021) explored the learning outcomes and flow status of learners in game-based learning. The results of the study found that learners not only had a higher flow status during the game but also had a greater understanding of what they were learning. Flow refers to the status in which a human being focuses on something and enjoys it (Csikszentmihalyi, 2000). Kiili et al. (2012) even mentioned in their study that when students learn through games, they are quite engaged, which can enhance their flow status.

Yang et al. (2020) mentioned that as the difficulty of the task set by the game mechanism increases, the level of flow status and anxiety will change accordingly. When the challenge is too difficult, it will cause anxiety. Therefore, different levels of game anxiety have different effects on learners in game-based learning. Lin and Hou (2022) found that learners' game anxiety was reduced through game-based learning. Ng et al. (2022) mentioned in their study that learning anxiety was significantly reduced when learners intensively utilized game-based learning.

In addition, role-play integrates well with game-based learning. In role-playing games, learners take on the role of their choice, experiencing the character's identity, undertaking tasks, and exploring (Miller et al., 2011). In addition, role-playing is interactive, and learners can gain experience by interacting with the characters or their peers in the game (Gartmeier et al., 2015; Stevens, 2015). Role-playing games not only allow learners to see things from a different perspective but even to reflect further (Grose-Fifer, 2017). Chen and Wu (2021) found that role-playing games enhanced learning motivation and even improved critical thinking. In addition, NPCs (Non-Player Characters) are common non-player-controlled characters in games, often responsible for providing information and clues to the player. Interactivity is one of the features of role-playing educational games where learners are able to acquire knowledge while interacting with others (Gartmeier et al., 2015). Role-playing educational games are often paired with NPCs

as a scaffolding mechanism, which allows learners to be more engaged in the simulated learning situation and to interact with the NPCs (Chien et al., 2023).

Instant feedback interactive scaffolding

The basis of the scaffolding theory comes primarily from the idea that learners are guided and supported in their learning (Wood et al., 1976). Rosenshine and Meister (1992) stated that scaffolding can be a useful strategy for facilitating the learning process by providing a supportive environment for the learner and at the same time reinforcing the learner's independence. Scaffolding also provides and encourages learners to interact more frequently with peers or teachers, enabling learners to engage deeply in learning activities (Sobhanzadeh et al., 2017). Moreover, building scaffolding can provide some feedback guidance for learners and also stimulate and develop their thinking and understanding (Makar et al., 2015). Also utilizing multiple scaffolding types of multidimensional scaffolding is helpful for increasing the effectiveness of digital game-based learning (Hou, 2022).

In educational games, NPCs can be a source of scaffolding provision. However, it is difficult to provide the same level of interaction as a real person. The use of real-person NPCs as scaffolding can provide learners with more realistic dialogues and interactions (Liu et al., 2021), reduce activity anxiety (Chan et al., in press), and enhance the flow status of learning (Chien et al., 2022). Chan et al. (in press) found that real-person NPCs with multidimensional scaffolding guidance can provide clear task objectives and achieve learning goals through instant feedback interactive experiences. However, real-person NPCs scaffolding is relatively costly in terms of activity costs as the instructor plays the role in the game (Chan et al., in press). To solve this problem, this study proposed the use of generative AI scaffolding to replace the scaffolding orientation of real-person NPCs' interactive scaffolding to achieve bidirectional and simulated pulsed dialogues. Being accompanied by an interactive peer or advisor as a dynamic instant feedback scaffolding can reduce anxiety (Rahman & Watanobe, 2023). However, *ChatGPT* is actually just a text generator, and at the moment it is still more difficult to get an accurate picture of the information it generates and to determine whether it is accurately related to the content of the question asked (Mhlanga, 2023). Therefore, exploring how to use the advantages of generative AI to develop learners' ability to retrieve is a topic worth investigating (Sánchez-Ruiz et al., 2023). The topic of generative AI as scaffolding in educational games has been less explored. This study investigated the advantages and limitations of generative AI as instant feedback scaffolding for learning, as well as the learner's experience and multiple psychological responses (e.g., flow status, motivation, and anxiety) during AI prompting in interactive scaffolding.

Scaffolding based contextualized game-based learning

In a game-based learning environment, learners can be guided to achieve learning goals through the natural progression of instructional content in the game's storyline (Wouters & van Oostendorp, 2016). When characterization is combined with narrative storytelling embedded in the game, it allows learners to take control of their virtual presence in the learning environment, satisfying their need for self-expression and increasing their learning motivation (Alexiou & Schippers, 2018). Scaffolding in contextualized game-based learning focuses on three aspects of the play process: interaction, narrative, and knowledge (Zuo et al., 2022). As mentioned above, generative AI has the potential to provide dynamic and interactive instant feedback, so for this study, we designed an online contextualized educational game with generative AI scaffolding to assist learning.

ChatGPT message generation can provide daily life interaction, generating human-like natural linguistic responses to a user's questions, and it is generated in a coherent and systematic way (Lo, 2023). This technology provides educators and learners with more diverse modes of interaction, including personalized learning support for learners (Zawacki-Richter et al., 2019),

accessible content construction (Pavlik, 2023), and pulsed interactive conversations (Terwiesch, 2023). *ChatGPT* can generate structured problem-solving feedback (Sánchez-Ruiz et al., 2023). Correct or partially correct feedback and suggestions can also be obtained based on the problem description, contextualization, and question and answer style (Rahman & Watanobe, 2023). These features are particularly well suited for providing interactive and realistic dialogues in situational games, which serve as instant feedback scaffolding that allows learners to consult guides and explore knowledge at any time.

Research objectives and questions

To summarize the above, the researchers designed a role-playing educational game and a contextual digital educational game with generative AI (using the *ChatPDF* service based on *ChatGPT* API) as the interactive scaffolding with a focus data source. The main learning goal was to promote learners' AI prompting skills and allow them to explore the scientific knowledge and spatial perception required to crack each level in the game. Generative AI employs a dialog-based approach that allows users to pose queries in natural language. Currently, the use of generative AI still has potential drawbacks, such as the possibility of errors or misleading information appearing in generative AI due to the unclear structure or semantics of the prompts by the user. Therefore, this study aimed to understand whether the content and history of users' conversations with generative AI have high accuracy and relevance of retrieval results guided by the cognitive scaffolding and to explore user feedback on the usefulness of the AI scaffolding.

This study also analyzed and evaluated several important psychological factors for learners in situational game-based learning. Chang et al. (2017) mentioned in their study that flow status can reflect the level of learner immersion and sense of engagement in educational game-based learning. Coffland and Huff (2022) also mentioned that appropriate gamification mechanisms and technologically presented digital tools can reduce students' anxiety about learning. Therefore, learners' flow status and anxiety are important analyzing factors in game-based learning (Hung et al., 2014), including learners' flow status during the situational game context (Zuo et al., 2022) and learners' possible anxiety during the game (Spieler et al., 2020). In addition, game-based learning is provided by integrating game elements with the learning goals and achieving learning by putting learners in control of the game and interacting with them (Ros et al., 2020). We addressed learners' feedback on the game (including usefulness, ease of use, and perception of game elements) and on the usefulness of the various types of scaffolding provided in the game, as well as the content of the learner-AI interaction.

This study adopted a single-group quasi-experimental design. The research questions were as follows:

1. What were the learners' flow status, game anxiety, scaffolding usefulness, and game feedback during the game tasks?
2. How many times did learners use *ChatPDF* to interact with the game, and what was the correct rate of generation of scaffolding guidance during the game tasks?
3. What was the correlation between flow status, anxiety, scaffolding usefulness, game feedback, game performance, and the number of times *ChatPDF* was used in a game task?

Online role-playing educational game with generative AI interactive scaffolding

Game mission and process

In this study, basic science knowledge was used as the teaching content to design the game, "*The Unfinished Cave Mission*." Set in 2095 A.D., aliens invade Earth, seize classified documents, and hide them in underground caves. Colonel Jack of the MEG Corps was the first one to go

into the caves to carry out the mission of retrieving the classified documents, but he unfortunately died in the course of duty. However, a memory chip was implanted in his brain. This chip had recorded all the memories of his lifetime, and the chip data were stored in the cloud data area in the format of AI, including all memories related to the execution of tasks. Players take the role of warriors in the game, and in the course of the adventure, they use the *ChatPDF* retrieval tool to ask questions to explore the memory information in Colonel Jack's brain, to avoid the dangers, and find the key clues to complete the mission. Since memories may have imprecise or fragmented characteristics, similar to the limited accuracy of ChatGPT's current generation technology, we used *ChatPDF* to simulate the memory dialogue between the players and Jack, and let the players practice the AI prompting skill (the skill of getting the required information in a generative AI system by giving appropriate prompts). The game tests the players' basic scientific knowledge, spatial perception ability, and AI information searching skills.

This study used *ChatPDF* (https://www.chatpdf.com/) as a chatbot with a chip for storing human memory. Considering that *ChatGPT* generally has a wider scope of questions and answers and is prone to generating more inaccurate information, *ChatPDF* limits the scope of the generated feedback mainly to a PDF file, so that when the designer writes the relevant story information in the PDF file, the player can get the clues of the game through AI searching. (In this study, an earlier version of *ChatPDF* was used, and learners were only able to retrieve and not see the PDF file). If the users' prompts are more accurate, *ChatPDF* may provide the answers the learners need. Each player can only ask a maximum of 50 prompts, and the game also uses *Google Forms* as a platform, allowing players to carry out tasks in the game. The player has to read about the cave scenes and events described in the form and choose the appropriate action in the form. Different action choices will take them to different episodes and different events, ultimately leading to a different ending. As shown in Figure 1, the left side is the adventure game form (*Google Form*), and the right side is the player opening *ChatPDF* to conduct clue exploration in another sub-page.

Before the game began, instructions and guidance were provided by the researcher. At the same time, an "Exploration Strategy Tips" note (Cognitive Scaffolding) was provided in the form of a picture that players can view at any time during the activity to review and reflect on their exploration strategies. The content included "AI Chip Search Strategy Hints" and "Spatial Exploration Strategy Tips." AI Chip Search Strategy Hints were used to remind players of some

Figure 1. Game screenshot of "The Unfinished Cave Mission."

prompt strategies during the searching process, such as "Avoid using a single keyword; it is recommended to use complete questions to make inquiries"; the Spatial Exploration Strategy Tips reminded players of the game strategies for orientation positioning, such as "Pay attention to the relationship between absolute orientation (east, south, north, and west) and relative directions (front, back, left, and right)."

Players used *Google Form* to play an adventure game where they must make the right choices and take the right actions. At the same time, they must also make full use of generative AI prompting techniques to retrieve key information, such as the path Colonel Jack had taken, the things he had done, and the dangers he had encountered from the memory chip (*ChatPDF*) implanted in his brain to help the player perform the unfinished cave tasks. There are a total of 20 levels in this game. Players must complete the level within the time limit, and those who do not complete the level within the set time are not allowed to do it again. In addition, if unsuccessful, the player can choose to go back to the previous decision point, re-select a different solution path, and then continue. The algorithm to calculate the learner's game performance is 5 points for passing each level, with 1 point deducted for each return action. For example, if a player succeeds within the time limit (completing all 20 levels) and returns 15 times, the learning performance score is $5 \times 20 - 15 = 85$.

Scaffolding mechanism in the game

"*The Unfinished Cave Mission*" is a scaffolding-based educational game. There are two types of scaffolding mechanisms in the game: instant feedback cognitive scaffolding and strategic scaffolding (as shown in Table 1). The multi-dimensional scaffolding structure that uses more than one type of scaffolding in educational games has been discussed by many studies that have found that it has a positive effect on learning (Hou, 2022; Hou et al., 2021).

This study's cognitive scaffolding was characterized by the "Instant Feedback Interactive Scaffolding" provided through *ChatPDF*, which is different from the general cognitive scaffolding of one-way giving. The instant feedback interactive scaffolding has two-way dynamics and explorability, and the player can obtain scientific knowledge and task clues through questions and answers with AI. In addition, the strategy scaffolding is provided by the "Mission Strategy Hints." The content includes "AI Chip Search Strategy Hints" and "Spatial Exploration Strategy Tips" to provide the strategies and suggestions related to AI prompting and spatial exploration in the mission, which guides the player to think strategically.

Methods

Participants

In this study, we recruited Taiwanese college and university students over 20 years old who were interested in this game as participants through an open online recruitment method. There were

Table 1. Two types of scaffolding in the game.

Scaffolding category	Scaffolding mechanism	Contents
Instant feedback interactive scaffolding	Interaction with *ChatPDF*	In the exploration stage of the game, scientific knowledge, various clues, guidance, and prompt hints that players can obtain during the dialogue with *ChatPDF*, including various events that have occurred in game tasks, scientific knowledge, spatial paths, clues to solve problems, etc., provide cognitive scaffolding that facilitates cognitive thinking.
Strategic scaffolding	Mission strategy hints	The prompts provide strategy hints for AI prompting and spatial exploration, promoting learners to reflect on game strategies at any time as strategic scaffolding.

59 participants (32 females and 27 males) aged 20 years or above. Some had experience with generative AI (72.88% had used *ChatGPT)*, while the others had no experience, but none of them had ever used generative AI in a game-based learning activity as scaffolding. This study was conducted using a single group to investigate the psychological factors and behavioral processes of players using generative AI, and all learners voluntarily participated by completing the informed consent form before the activity. The consent form included the experimental content, experimental procedures, and participants' rights. The research ethics procedures and informed consent for this study were reviewed by the Office of Research Ethics, National Chengchi University, Case No. (NCCU-REC-202112-E111).

Procedure

The experimental procedure of this study was as follows: Before the beginning of the activity, the researcher explained the procedure of the activity (10 min), and then the players were asked to perform the game task (40 min; if it took longer than 40 min, they failed the task). After the game was completed, the participants were given 20 min to fill out the questionnaires on flow, activity anxiety, scaffolding usefulness, and game feedback.

Research tools

Flow scale

For measuring the learners' flow status during learning activities, Kiili's (2006) flow scale was used as a flow measure; it was translated into Chinese by Hou and Li (2014). This study used this translated version for testing. There were two dimensions, Flow Antecedents and Flow Experience, with a total of 22 questions. Flow Antecedents were divided into five sub-dimensions, namely Challenge-Skills Balance, Clear Goals, Unambiguous Feedback, Sense of Control, and Action Awareness Merging, while Flow Experience was divided into four sub-dimensions, namely Concentration, Time Distortion, Autotelic Experience, and Loss of Self-Awareness. This questionnaire used a 5-point Likert scale. After analyzing the data of this study sample, the overall reliability of the Flow Scale (Cronbach's $\alpha = .90$) was extremely high.

Activity anxiety scale

To explore the level of anxiety of the learners while participating in the activities, this study referred to the Affective Filter Hypothesis developed by Higgs and Krashen (1982), and the anxiety-related questions Chinese version of the Learning Experience Scale adapted by Hung (2001) were used to assess the learners' anxiety level; it consisted of eight questions. This study directly used this Chinese version of the questionnaire, which used a 5-point Likert scale. After analyzing the data of this study sample, the overall reliability of the Activity Anxiety Scale (Cronbach's $\alpha = .87$) was high.

Game feedback scale

To explore the learners' perceptions of the use of the game, this study modified the technology acceptance scale proposed by Davis (1989). Modifying and increasing the questions on learners' perceptions of the game elements of the game itself made it more consistent with the content of this study. The entire game feedback questionnaire included three dimensions: usefulness, ease of use, and game element perception. Usefulness explored learners' perceptions of whether the game contributed to knowledge learning after participating in the activity. Ease of use explored whether the game was easy for learners to understand and operate after participating in the activity. In addition, the assessment of game elements was conducted with reference to the elements that promote game motivation as proposed by Hou (2016), with a total of 12

questions. This questionnaire used a 5-point Likert scale. After analyzing the data of this study sample, the overall reliability of the Game Feedback Scale (Cronbach's $\alpha = .81$) was high.

Scaffolding Usefulness Questionnaire

This study explored learners' feedback on the usefulness of various types of scaffolding provided in the game for learning. This study adapted the Scaffolding Usefulness Questionnaire developed by Hou et al. (2021), and the categories of scaffolding were modified based on the content of this study. The scaffolding usefulness questionnaire was divided into two dimensions: "strategic scaffolding" and "instant feedback interactive cognitive scaffolding" to separately investigate learners' perceptions of the various effects of these two types of scaffolds on learning. This questionnaire used a 5-point Likert scale. After analyzing the data of this study sample, the overall reliability of the Scaffolding Usefulness Questionnaire (Cronbach's $\alpha = .86$) was high.

Gaming behavior coding scheme

In this study, to investigate the content analysis of learners' interaction with *ChatPDF* in the game "*The Unfinished Cave Mission*," this study modified Chan et al. (in press) Game Behavioral Coding Sheet to design the scheme. There were two aspects in the coding table, Player Question (Q) and *ChatPDF* Answer (A), with a total of 13 codes, as shown in Table 2.

Data analysis tools and methods

IBM SPSS Statistics version 26 was used in this study to analyze the data, and the method of data analysis for each research question is described as follows:

1. Flow status, activity anxiety, game feedback, and scaffolding usefulness were analyzed using a one-sample *t*-test and were compared to the median of the scale (i.e., 3). This was used to support the descriptive statistical analysis and to see if each variable was significantly higher than 3.
2. As for interactive behavioral analysis, the coding scheme has 13 codes. The coders each performed a complete coding of the same group of data and calculated Cohen's Kappa coefficient, which in this study was 0.89, achieving a high degree of inter-rater agreement. The number and proportion of times and proportions were analyzed by descriptive statistics for all groups of coded data at a later stage.
3. The correlations among flow status, activity anxiety, scaffolding usefulness, game feedback, learning performance, and the number of times *ChatPDF* was used were analyzed using Pearson correlation analysis.

Results and discussion

Flow status and activity anxiety

In this study, a one sample *t*-test (compared to 3, the median of the scale) was conducted on the flow status and activity anxiety to see if the degree of flow status and activity anxiety was above 3. In terms of flow status, as shown in Table 3, it was found that the overall average score of flow, as well as the scores of each sub-dimension of flow antecedents, and flow experience were significantly higher than 3, and most of the sub-dimensions even had mean scores higher than 4. Among them, on the sub-dimension of flow experience, learners' *Concentration* score was higher ($M = 4.53$), from which it could be inferred that learners were fully engaged in the game (Buchanan & Csikszentmihalyi, 1990; Perttula et al., 2017). In addition, learners' *Autotelic Experience* was high ($M = 4.49$), indicating that learners could focus their attention on

Table 2. Gaming behavior coding scheme.

Type	Code	Behavior	Description
Players ask questions (Q)	KQ	Knowledge content questions	Learners asked the AI about knowledge related to science. For example, "Which chemical can be used to dehydrate a leech?"
	PQ	Puzzle clue questions	Learners asked the AI for puzzle clue information. For example, "What dangers did Jack encounter in which numbered cave?"
	SQ	Spatial clue questions	Learners asked the AI about spatial clues. For example, "Which way did he choose? Where is the leech pond?"
	OQ	Other clue questions	Learners asked the AI for other clues. For example, "How did Jack die in the deep well of darkness?"
	IQ	Incomplete or irrelevant questions	Learners asked the AI incomplete or irrelevant questions. For example, "I would like to ask"
ChatPDF answer (A)	A1	Correctly provide knowledge content information	AI provided learners with the right knowledge related to science. For example, "The chemical formula for salt usually refers to sodium chloride (NaCl). It is a crystalline solid composed of one sodium ion and one chloride ion. Its chemical formula is NaCl."
	A2	Correctly provide puzzle information	AI provided learners with the correct clue information for puzzles. For example, "The cave number that Jack saw when he fell was HX252, which should be the number of the cave."
	A3	Correctly provide spatial clue information	AI provided learners with the correct orientation route information. For example, "According to the information provided, Jack's path before he lost his footing was as follows: facing east in Totem's cave, he turned right at the first intersection, then right again at the next intersection, then left at the next intersection, and left again at the last intersection. At that point he was still facing east and reached the platform of the secret vault with the door."
	A4	Correctly provide other clue information	AI provided learners with the correct information on other clues. For example, "Jack encountered a variety of dangers in the cave, including flaming corridors, stone door traps, and deep wells. He was also attacked by R guards and eventually died from his injuries when he fell into a deep well. In the caves, Jack needed to take care of the various dangers in order to complete his mission."
	A5	No corresponding information was available, but other indirect advice or knowledge was given.	AI did not provide the learner with a corresponding answer but gave other advice or knowledge. For example, "There is no mention of whether Jack used the solution in the yellow cylinder. It is recommended to know the nature and purpose of the solution before use to ensure safety."
	A6	No corresponding information was provided, and no other recommendations were made	AI provided no information to learners. For example, "From the information provided, there is no mention of whether the site where Jack lost his footing is located in the direction of the site infected by the microorganisms. Therefore, it is not possible to answer this question."

the game and feel satisfied. Huang and Hew (2018) stated that enjoyment and participation can be maximized when optimal mental and physical states are felt during activities.

In terms of anxiety, the overall anxiety score for learning was significantly lower than 3 ($M = 2.44$, $p < .001$), indicating that learners were not particularly anxious about the way the game was conducted. Reed and Ferdig (2021) found that interactive games are effective in terms of reducing learning anxiety. In addition, Pallavicini et al. (2021) suggested that simulation-based educational games can reduce learner anxiety if they are designed with good situation simulation.

Table 3. Descriptive statistics and one-sample *t*-test results of the flow dimensions and activity anxiety.

	(*n* = 59)			
	M	*SD*	*t*	*p*
Overall flow	4.24	0.46	20.53***	.000
Flow antecedents	4.16	0.56	16.09***	.000
Challenge-skills balance	4.23	0.58	16.21***	.000
Clear goals	4.35	0.69	15.00***	.000
Unambiguous feedback	3.90	0.73	9.46***	.000
Sense of control	4.25	0.75	12.83***	.000
Action-awareness merging	4.08	0.78	10.64***	.000
Flow experience	4.31	0.47	21.33***	.000
Concentration	4.53	0.59	20.07***	.000
Time distortion	4.10	0.80	10.54***	.000
Autotelic experience	4.49	0.59	19.45***	.000
Loss of self-consciousness	3.70	0.96	5.65***	.000
Overall activity anxiety	2.44	0.61	7.04***	.000

***$p < .001$.

Table 4. Descriptive statistics and one-sample *t*-test results of the game feedback.

	(*n* = 59)				
	M	*SD*	*t*	*p*	*ES*
Game feedback overall score	4.26	0.61	15.81***	.000	2.06
Usefulness overall score	4.25	0.68	14.19***	.000	1.85
Usefulness for knowledge of concepts related to the application of science to life	4.15	0.78	11.30***	.000	1.47
Usefulness for conceptual knowledge of AI prompting	4.12	0.87	9.85***	.000	1.28
Usefulness for AI prompting	4.49	0.77	14.80***	.000	1.93
Ease of use overall score	4.26	0.74	13.16***	.000	1.71
Easy to operate	4.37	0.81	13.06***	.000	1.70
The logic of the plot is easy to understand	4.29	0.79	12.54***	.000	1.63
No problems with the operation	4.12	1.07	8.05***	.000	1.05

***$p < .001$.

Game feedback

In this study, a one sample *t*-test was conducted on game feedback (compared to the median 3 of the scale); the results are shown in Table 4. Overall Game Feedback, Game Usefulness, and Game Ease of Use scores were all significantly higher than 3, indicating that learners felt that they gained knowledge from the game and the game was easy to play. In particular, in the *Game Usefulness* section, learners scored the highest on the item that they felt the game helped them learn about AI prompting ($M = 4.49$), which also indicated that learners felt that the game helped them develop their AI prompting skills. Lim et al. (2023) pointed out that the feedback results generated by generative AI are highly dependent on the prompt it receives. Therefore, the cultivation of learners' AI prompting ability was very important for generative AI to play an effective powerful role. In addition, Mogali (2023) raised concerns about ChatGPT's well-written but inaccurate messages, and Jalil et al. (2023) also mentioned the concern that "currently *ChatGPT* has poor judgment of its own accuracy." Therefore, it is important for researchers and learners to cultivate the knowledge related to judging messages, in addition to strengthening the AI prompting ability to retrieve them, to determine whether the generative AI can give correct information as feedback. The ease of use sub-dimension also showed that learners found the game using *ChatPDF* easy and smooth to operate, and the logic of the plot was easy to understand. Spieler et al. (2020) mentioned that poor mechanisms and rules of gamified learning may lead to students' cognitive load or learning anxiety. This also means

Table 5. Descriptive statistics and one-sample *t*-test results of the game elements.

	(*n* = 59)				
	M	*SD*	*t*	*p*	*ES*
Game elements	4.30	0.66	15.14***	.000	1.97
Sense of control	4.07	0.83	9.91***	.000	1.29
Uncertainty	4.39	0.72	14.84***	.000	1.93
Achievement	4.31	0.88	11.44***	.000	1.49
Think the game was fun	4.56	0.68	17.71***	.000	2.31
Wish to play again	4.17	0.91	9.84***	.000	1.28

***$p < .001$.

that the design mechanism of this study, which applied intuitive and realistic generative AI chat rooms and game interfaces, was suitable for game-based learning and might not increase external cognitive load.

In terms of game elements, the scores of all dimensions were significantly higher than 3. The results are shown in Table 5. The results showed that the learners thought the game was fun, and they had a high sense of control and high uncertainty in the game. Hassan et al. (2019) mentioned that good game elements could elicit learners' achievement, engagement, and elevated motivation. According to the study, an educational game with generative AI as scaffolding created a sense of achievement and a desire to play again for learners.

Scaffolding usefulness

In this study, a one sample *t*-test was conducted for the scaffolding usefulness (compared to the median 3 of the scale); the results are shown in Table 6. The scores of the usefulness of overall scaffolding, strategic scaffolding, and instant feedback interactive scaffolding of each sub-dimension were all significantly higher than 3. The "Mission Strategy Hints" was strategic scaffolding, and the results showed that the strategic scaffolding in the game was useful and helpful for AI prompting and spatial exploration in this game. In addition, the game provided instant feedback interactive scaffolding through interactive conversations with *ChatPDF*, and the results showed that it helped learners analyze the context of the game tasks, made them more motivated to play the game, and facilitated strategic thinking, and reduced anxiety during the game. Xun and Land (2004) thought that interaction during activities could act as a kind of soft scaffolding. Diverse modes of interaction, on the other hand, could lead to more feedback from learners (Urmeneta & Walsh, 2017). It has been found that a good scaffolding support combined with a distance gamification mechanism does not bring more anxiety to the learners due to the distance education context, and allows students to have high engagement and positive emotions (Chen et al., 2023). This study proposed the use of generative AI as instant feedback. Interactive scaffolding was cognitive scaffolding that provided a dynamic and interactive conversation context instantly, which was suitable for helping learners to search for data, analyze contexts, and think about strategies instantly during the learning activity, and then enhance the effectiveness of learning. This type of scaffolding was different from the common conceptual scaffolding that provided fixed knowledge and guidance and had the potential to be a much more useful aid to learning.

Correlation analysis

In this section, Pearson correlation analysis was conducted on learners' flow status, activity anxiety, game elements, game usefulness, ease of use, scaffolding usefulness, and game learning performance. *ChatPDF* usage times were used to find out whether there was a correlation between these dimensions. The results are shown in Table 7.

Table 6. Descriptive statistics and one-sample *t*-test results of scaffolding usefulness.

	(*n* = 59)				
	M	*SD*	*t*	*p*	*ES*
Overall scaffolding	4.15	0.66	13.37***	.000	1.74
Strategic scaffolding (mission strategy tips)					
Helped me understand more about how to retrieve effective key words and phrases	4.19	0.82	11.12***	.000	1.45
Helped me to be more aware of the use of absolute and relative positions	3.86	1.07	6.18***	.000	0.80
Instant feedback interactive scaffolding (interactive conversations with *ChatPDF*)					
Helped me learn more about science knowledge in the game	4.02	0.97	8.02***	.000	1.04
Helped me analyze the game's mission context	4.42	0.67	16.21***	.000	2.11
Helped me think strategically during the game	4.34	0.82	12.51***	.000	1.63
Helped me to be more motivated to play	4.37	0.67	15.81***	.000	2.06
Helped reduce my anxiety in the game	3.86	1.17	5.69***	.000	0.74

***$p < .001$.

Table 7. Pearson correlation analysis.

	Flow	Activity anxiety	Game elements	Game usefulness	Game ease of use	Scaffolding usefulness	Game performance	ChatPDF usage times
Flow	1.00							
Activity anxiety	−0.28*	1.00						
Game elements	.74***	−0.13	1.00					
Game usefulness	.56***	.05	.66***	1.00				
Game ease of use	.65***	−0.27*	.67***	.49***	1.00			
Scaffolding usefulness	.71***	−0.15	.77***	.72***	.64***	1.00		
Game performance	.27*	−0.25	.21	.15	.28*	.25	1.00	
ChatPDF usage times	.08	.08	−0.02	.17	−0.30*	.05	−0.26*	1.00

*$p < .05$, ***$p < .001$.

In terms of flow status, this study found a significant negative association between activity anxiety and flow status. Liu et al. (2011) mentioned that students who felt high anxiety during a simulation game did not utilize the learning strategies provided in the game as smoothly as students in high flow status. Lin and Hou (2022) mentioned that the cognitive scaffolding provided in educational games may reduce learners' activity anxiety. The "Instant Feedback Interactive Scaffolding" in this study might enable learners to interact with *ChatPDF* in a two-way manner. However, when incorrect retrieval information was generated too often, it might increase the learners' cognitive load and increase their activity anxiety, and might cause flow status discontinuity. In addition, flow status was significantly and positively related to game elements, game usefulness, game ease of use, scaffolding usefulness, and game learning performance. In this game, flow status was associated with almost all the other variables, so it could be inferred that flow status is an important indicator of game-based learning, which is the same as was pointed out by Bressler and Bodzin (2013) study. It is further inferred that when activity anxiety is lower, the learner may maintain a high flow, and may be more involved in the game, which is the same as Hou's (2015) research pointed out.

In addition, this study found that game ease of use was negatively related to activity anxiety, which could be attributed to the fact that when the game operation was not smooth, or learners were not sure about *ChatPDF* responses, it would increase the learners' cognitive load and cause them anxiety. This was like Yang et al. (2020) finding that anxiety is induced when the cognitive load is too high. Scaffolding usefulness has a significant positive correlation with game elements, game usefulness, and game ease of use. It is possible that the more useful the scaffolding is to

the learner, the better the design of the game elements, the more useful the game is to learners, and the easier the game is to play. There was a significant positive correlation between game effectiveness and game ease of use, and it was possible that when learners found one game easier to use, they would perform better. There was a significant negative correlation between *ChatPDF* usage times and the game's ease of use as well as game performance. It is possible that learners may have felt that the more times they asked a question without getting an answer, the more difficult the game was to use, the less easy it was to manipulate the game, and the worse their performance was. Therefore, higher *ChatPDF* usage times were not necessarily better, and it may be necessary to consider the learners' AI prompting ability.

Behavioral content analysis

This study coded the learners' *ChatPDF* interactions during the game, with the main purpose of exploring the learners' questioning and *ChatPDF* responses.

Regarding players asking questions (Q), a total of 1251 inquiries were made to *ChatPDF* by 59 learners. As shown in Table 8, learners used more queries to explore the puzzle clues section (56.99%) and only 2.16% were incomplete or irrelevant queries, which indicates that they overwhelmingly asked appropriate questions and mainly explored puzzle clues when asking questions using generative AI.

Next, we discuss the proportions of codes of the various responses of *ChatPDF* for the different types of questions, as shown in Table 9.

The knowledge content questions (KQ)

ChatPDF provided 48.57% correct answers (A1), 29.71% indirect suggestions (A5), and 21.71% irrelevant responses (A6). It showed that when learners made knowledge-related inquiries every time, the percentage of relevant answers and suggestions (A1 and A5) was 78.28%. Among them, nearly 50% of the questions were answered correctly. Since players can have multiple opportunities to ask questions to modify the structure of the entire question and the direction of the

Table 8. Statistics on the content classification and frequency of questions asked by the 59 learners.

Question type	Number of times	Percentage
Knowledge content questions (KQ)	175	13.99%
Puzzle clue questions (PQ)	713	56.99%
Spatial clue questions (SQ)	234	18.71%
Other clue questions (OQ)	102	8.15%
Incomplete or irrelevant questions (IQ)	27	2.16%
Total number of questions	1251	100.00%

Table 9. Correspondence between players asking questions (Q) and AI answers (A).

Codes of players asking questions (Q)	Codes of AI answering (A)	Percentage
Knowledge content questions (KQ)	Direct answer (A1)	48.57%
	Indirect suggestions (A5)	29.71%
	Irrelevant responses (A6)	21.71%
Puzzle clue questions (PQ)	Direct answer (A2)	58.63%
	Indirect suggestions (A5)	24.96%
	Irrelevant responses (A6)	16.41%
Spatial clue questions (SQ)	Direct answer (A3)	49.15%
	Indirect suggestions (A5)	41.45%
	Irrelevant responses (A6)	9.40%
Other clue questions (OQ)	Direct answer (A4)	67.65%
	Indirect suggestions (A5)	15.69%
	Irrelevant responses (A6)	16.67%
Incomplete or irrelevant questions (IQ)	Indirect suggestions (A5)	55.56%
	Irrelevant responses (A6)	44.44%

question, they could increase the chance of obtaining clues that are beneficial to solving the task. Therefore, when learners asked questions in the knowledge content part, the proportion of directly obtaining correct answers should be high. However, about 20% of them were unable to generate information useful for knowledge learning.

The puzzle clue questions (PQ)

ChatPDF provided 58.63% correct answers (A2), 24.96% indirect suggestions (A5), and 16.41% irrelevant responses (A6). It means that each time a learner asked a question related to the puzzle clue, 83.59% of them got relevant answers and suggestions (A2 and A5). Thus, more than 80% of the learners got the information related to the solution of the puzzle. Among them, the proportion of correct answers was more than twice as high as the proportion of indirect suggestions. It might be that learners could get the correct answer with fewer correction tips when asking questions in the clue part of the puzzle, and the proportion of directly obtaining the correct answer was higher. Learners seemed to be more precise in the question structure and question direction when asking puzzle clues, and *ChatPDF* answers were also more correct and accurate.

The spatial clue questions (SQ)

ChatPDF provided 49.15% correct answers (A3), 41.45% indirect suggestions (A5), and 9.40% irrelevant responses (A6). This means that when learners asked questions related to spatial cues, a high proportion (90.60%) of *ChatPDF*'s answers could allow learners to get relevant or suggested responses (A3 and A5), among which the proportion of correct answers and the proportion of indirect suggestions is close. It is worth noting that <10% of learners failed to obtain useful information when asking for spatial clues.

The other clue questions (OQ)

ChatPDF provided 67.64% correct answers (A4), 15.69% indirect suggestions (A5), and 16.67% irrelevant responses (A6). It means that when learners asked for other thread clues, the percentage of those who got relevant answers and suggestions (A4 and A5) was 83.34%. Among them, the proportion of correct answers was more than four times the proportion of indirect suggestions. Learners who asked questions for other clues could get the correct answer with fewer corrections, and the proportion of directly obtaining the correct answer was relatively high. When learners asked for other clues, the question structure and question direction could make *ChatPDF* answers more accurate. In addition, the inability to obtain useful information was <20%.

The incomplete or irrelevant questions (IQ)

ChatPDF provided 55.56% indirect suggestions (A5), and 44.44% irrelevant responses (A6). Even in the face of incomplete questions, AI still provided alternative suggestions or information in more than 50% of cases, but more than 40% were unable to obtain useful information.

In addition, the proportion of *ChatPDF* providing direct answers or related answers and suggestions in this game totaled 83.54%, as shown in Table 10. Jalil et al. (2023) explored the correctness of *ChatGPT* responses in a software testing session. According to the results of the study, *ChatGPT* was able to provide correct or partial answers to 55.6% of the questions. In

Table 10. ChatPDF direct answer or related answer and suggestion ratio total.

Codes of *ChatPDF* answering	Percentage
Direct answer	54.92%
Indirect suggestions	28.62%
Irrelevant responses	16.47%

this study, when learners asked for information, an average of more than 80% received direct or indirect responses, and even if the correct information could not be provided immediately, a certain percentage of them still gave advice or other information to the learners. Among them, statistical analysis shows that the proportion of correct answers was about twice that of indirect suggestions. It showed that the use of tools that can focus on specific files as the source of generation (e.g., *ChatPDF*) should have more potential to help guide learning. In addition, according to the results of the analysis of the question "I think the game can help me understand more relevant knowledge for AI prompting" in the above-mentioned game feedback questionnaire ($M = 4.49$), it can be seen that the scaffolding in the game might be able to guide learners in the question structure and question direction to achieve better prompting. Also, studies have found ChatGPT's inaccurate message responses to be worrisome (Mogali, 2023). This study found that there were still 10–20% of AI-generated responses that failed to respond to the question or which may have been inaccurate. Therefore, in addition to strengthening the ability of AI prompting, the ability of learners to determine the correctness of messages was also very important. The accuracy of *ChatGPT* responses depended on several factors, including the specificity of the prompt content, the complexity of the prompt or topic, and the need to consider the domain of knowledge and the specific context of use (Wardat et al., 2023). This also echoed the findings of this study that the percentage of correct or relevant responses varies with the type of prompts asked by learners and their prompting skills.

Conclusion and recommendations

Conclusion

To summarize the above analysis, this study designed a role-playing educational game using generative AI as "instant feedback interactive scaffolding," which can provide interactive scaffolding with a human-to-human-like communication context through the instant dynamic feedback characteristics of generative AI. The findings of several research questions are discussed below, along with possible causes.

Flow, anxiety, scaffolding usefulness, and game feedback

First, this study found that the interaction with *ChatPDF* in this exploratory game could help learners achieve deeper contextual analysis, reduce their learning anxiety, and maintain a high level of mind flow. This finding is similar to previous studies' findings that when the provided scaffolding has a positive impact on learners' understanding of learning concepts, it will reduce their anxiety, which in turn will increase their learning satisfaction (Sobhanzadeh et al., 2017). Chien et al. (2023) research also pointed out that learners can adjust their exploration pace based on their understanding of the information context and improve their sense of self-learning control. Learners may feel the usefulness of instant feedback interactive scaffolding during the exploration process in the game.

Numbers of interactive behaviors and scaffolding guidance

In addition, after analyzing the accuracy rate of the responses provided by *ChatPDF* as scaffolding in the game task, the behavioral analysis shows that the accuracy rate of *ChatPDF*'s response is a higher percentage of direct answers or indirect suggestions (78.28% for knowledge content questions, 83.59% for puzzle clue questions, 90.6% for spatial clue questions, and 83.33% for other clue questions). In another study, Jalil et al. (2023) found that the general *ChatGPT* was able to provide correct or partial answers to 55.6% of the questions, whereas the generative AI in the present study used a generative scheme based on specific text file sources, and so seemed to be more accurate in providing responses. Therefore, such solutions (e.g., *ChatPDF*) may be suitable for instant feedback interactive scaffolding. Learners can continuously adjust

the content of their own prompts according to the responses of the generative AI, which enhances the sense of learning control and the usefulness of the scaffolding.

Correlation analysis

Last, the results of the correlation analysis showed that there was a significant negative correlation between flow and activity anxiety, and a significant positive correlation between flow and game elements, game usefulness, game ease of use, scaffolding usefulness, and game learning performance. This suggests that flow is a key psychological variable in this type of educational game experience. In this game, flow is associated with almost all dimensions, so it can be inferred that it is an important indicator of game-based learning, which is the same as what Bressler and Bodzin (2013) pointed out. Other studies have also suggested that contextual cues with perceived usefulness and ease of use can lead to sustained flow performance, which can be revealed by the player's flow experience and satisfaction in situational games (Liu et al., 2018).

Recommendations

Based on the above results, this study provides some recommendations as follows:

The use of generative AI in education for game-based learning activity design has been hotly debated and researched in recent years. Zhai (2022) pointed out that learners should be able to use generative AI tools to perform subject area tasks, and teachers should design AI-related learning tasks. Therefore, it is suggested that generative AI can provide instant feedback interactive scaffolding that can be used as learning agents in the game to assist learning by referring to the design mechanism proposed in this study. In addition, this study only adopted the form of a picture and text adventure game. In the future, it is suggested that virtual space can be added to the game as a simulation situation scenario design so that learners can experience a more complete immersive learning experience. Pears et al. (2020) pointed out that the use of experiential and interactive learning contexts to simulate the process of dealing with the details of the real case is crucial for promoting students' learning performance and is expected to achieve a higher degree of learning transfer.

As for generative AI as scaffolding, Sallam (2023) pointed out that the use of *ChatGPT* in the educational domain may generate doubts related to its accuracy and reliability due to its limitations. The tool used in this study is based on the uploaded PDF content database as the main source of generation. As analyzed in this study, the generated responses given had a certain percentage (78 ~ 90%) of providing clues or information that was helpful for learning, and the learners also perceived the usefulness of their information. This type of AI tool with limited data sources should have more potential to be utilized as a scaffold for learning. Therefore, it is suggested that teachers can design scaffolding with *ChatPDF* interaction as supplementary scaffolding in teaching. However, this study found that about 10–20% of the responses still failed to obtain useful information or were not precise enough, so it is recommended that teachers still need to cultivate learners' ability to diagnose and interpret the information generated in the process of learning with AI.

Research limitations

The main focus of this research was to initially explore the psychological factors and behavioral processes of players using generative AI. It does not prove that this teaching method can have better learning effectiveness than other teaching methods. In the future, further comparisons can be made with other teaching methods (such as didactic instruction and other scaffolding delivery methods). In addition, considering the varying experience and basic abilities of learners in AI retrieval, it is recommended that there be a basic introductory course on the use of GPT

before starting the game. In addition, this research is based on specific plots and scientific knowledge, and the content and length of time used for retrieval were limited. Therefore, research in other different fields may have different accuracy rates.

IRB statement

The research process and scales of this study were reviewed by the Office of Research Ethics, National Chengchi University, Case No. (NCCU-REC-202112-E111).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Chih-Chung Chien is a PHD student of National Taiwan University of Science and Technology. His research interests include game-based learning and educational game design.

Hung-Yu Chan is a PHD student of National Taiwan University of Science and Technology. His research interests include game-based learning and educational game.

Huei-Tse Hou is a Distinguished Professor of National Taiwan University of Science and Technology, Taiwan. He is also the Director of Mini-Educational Game development Group in E-learning Research Center in National Taiwan University of Science and Technology (NTUST MEG). His research interest focuses behavioral pattern analysis and game-based learning.

Funding

This research was supported by the 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.

Data availability statement

The data that support the findings of this study are available from the corresponding author, upon reasonable request.

References

Alexiou, A., & Schippers, M. C. (2018). Digital game elements, user experience and learning: A conceptual framework. *Education and Information Technologies*, *23*(6), 2545–2567. https://doi.org/10.1007/s10639-018-9730-6

Bawa, P., Watson, S. L., & Watson, W. (2018). Motivation is a game: Massively multiplayer online games as agents of motivation in higher education. *Computers & Education*, *123*, 174–194. https://doi.org/10.1016/j.compedu.2018.05.004

Bressler, D. M., & Bodzin, A. M. (2013). A mixed methods assessment of students' flow experiences during a mobile augmented reality science game. *Journal of Computer Assisted Learning*, *29*(6), 505–517. https://doi.org/10.1111/jcal.12008

Buchanan, R., & Csikszentmihalyi, M. (1990). Flow: The psychology of optimal experience. *Design Issues*, *8*(1), 80. https://doi.org/10.2307/1511458

Chan, H. Y., Liu, S. W., & Hou, H. T. (in press). Interacting with real-person non-player characters to learn history: Development and playing behavior pattern analysis of a remote scaffolding-based situated educational game. *Interactive Learning Environments*. https://doi.org/10.1080/10494820.2023.2192745

Chang, C.-C., Liang, C., Chou, P.-N., & Lin, G.-Y. (2017). Is game-based learning better in flow experience and various types of cognitive load than non-game-based learning? Perspective from multimedia and media richness. *Computers in Human Behavior*, *71*, 218–227. https://doi.org/10.1016/j.chb.2017.01.031

Chen, H.-L., & Wu, C.-T. (2021). A digital role-playing game for learning: Effects on critical thinking and motivation. *Interactive Learning Environments*, *31*(5), 3018–3030. https://doi.org/10.1080/10494820.2021.1916765

Chen, M.-F., Chen, Y.-C., Zuo, P.-Y., & Hou, H.-T. (2023). Design and evaluation of a remote synchronous gamified mathematics teaching activity that integrates multi-representational scaffolding and a mind tool for gamified learning. *Education and Information Technologies*, *28*(10), 1–27. https://doi.org/10.1007/s10639-023-11708-6

Chien, C. C., Ho, Y.-T., & Hou, H.-T. (2023). Integrating immersive scenes and interactive contextual clue scaffolding into decision-making analysis ability training game. *Journal of Educational Computing Research*, *62*(1), 376–405. https://doi.org/10.1177/07356331231205058

Chien, C. C., Huang, S. T., & Hou, H. T. (2022). Design and evaluation of a contextual distance management training game with real-person non-player character mechanism. In *Proceedings of the 14th Asian Conference on Education (ACE2022)*.

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*, *31*(5), 3285–3302. https://doi.org/10.1080/10494820.2021.1926287

Coffland, D., & Huff, T. (2022). Stats kwon do: A case study in instructional design, multimedia and gamification of instruction. *TechTrends*, *66*(6), 945–956. https://doi.org/10.1007/s11528-022-00793-y

Csikszentmihalyi, M. (2000). *Beyond boredom and anxiety*. Jossey-Bass.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319–340. https://doi.org/10.2307/249008

Gartmeier, M., Bauer, J., Fischer, M. R., Hoppe-Seyler, T., Karsten, G., Kiessling, C., Möller, G. E., Wiesbeck, A., & Prenzel, M. (2015). Fostering professional communication skills of future physicians and teachers: Effects of e-learning with video cases and role-play. *Instructional Science*, *43*(4), 443–462. https://doi.org/10.1007/s11251-014-9341-6

Grose-Fifer, J. (2017). Using role-play to enhance critical thinking about ethics in psychology.

Hassan, M. A., Habiba, U., Majeed, F., & Shoaib, M. (2019). Adaptive gamification in e-learning based on students' learning styles. *Interactive Learning Environments*, *29*(4), 545–565. https://doi.org/10.1080/10494820.2019.1588745

Higgs, T. V., & Krashen, S. D. (1982). Principles and practice in second language acquisition. *The Modern Language Journal*, *67*(2), 168–169. https://doi.org/10.2307/328293

Hou, H. T. (2016). *Game-based Learning.* Commonwealth Magazine Co., Ltd.

Hou, H. T., & Li, M. C. (2014). Evaluating multiple aspects of a digital educational problem-solving-based adventure game. *Computers in Human Behavior*, *30*, 29–38. https://doi.org/10.1016/j.chb.2013.07.052

Hou, H.-T. (2022). Augmented reality board game with multidimensional scaffolding mechanism: A potential new trend for effective organizational strategic planning training. *13,* 932328. https://doi.org/10.3389/fpsyg.2022.932328

Hou, H.-T., Fang, Y.-S., & Tang, J. T. (2021). Designing an alternate reality board game with augmented reality and multi-dimensional scaffolding for promoting spatial and logical ability. *Interactive Learning Environments*, *31*(7), 4346–4366. https://doi.org/10.1080/10494820.2021.1961810

Hou, H-T. (2015). Integrating cluster and sequential analysis to explore learners' flow and behavioral patterns in a simulation game with situated-learning context for science courses: A video-based process exploration. *Computers in Human Behavior*, *48*, 424–435. https://doi.org/10.1016/j.chb.2015.02.010

Huang, B., & Hew, K. F. (2018). Implementing a theory-driven gamification model in higher education flipped courses: Effects on out-of-class activity completion and quality of artifacts. *Computers & Education*, *125*, 254–272. https://doi.org/10.1016/j.compedu.2018.06.018

Hung, C.-M., Huang, I., & Hwang, G.-J. (2014). Effects of digital game-based learning on students' self-efficacy, motivation, anxiety, and achievements in learning mathematics. *Journal of Computers in Education*, *1*(2–3), 151–166. https://doi.org/10.1007/s40692-014-0008-8

Hung, M. H. (2001). *The effects of captioned, subtitled and non-captioned television videotapes on foreign language learning* [Master's thesis]. National Cheng Kung University Institute of Education. https://hdl.handle.net/11296/sx7ca2

Jalil, S., Rafi, S., LaToza, T. D., Moran, K., & Lam, W. (2023, April 1). ChatGPT and software testing education: Promises & perils. IEEE Xplore. https://doi.org/10.1109/ICSTW58534.2023.00078

Kiili, K. (2006). Evaluations of an experiential gaming model. *Human Technology*, *2*(2), 187–201. https://doi.org/10.17011/ht/urn.2006518

Kiili, K., de Freitas, S., Arnab, S., & Lainema, T. (2012). The design principles for flow experience in educational games. *Procedia Computer Science*, *15*, 78–91. https://doi.org/10.1016/j.procs.2012.10.060

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.-C., & Hou, H.-T. (2022). The evaluation of a scaffolding-based augmented reality educational board game with competition-oriented and collaboration-oriented mechanisms: Differences analysis of learning effectiveness, motivation, flow, and anxiety. *Interactive Learning Environments*, *32*(2), 502–521. https://doi.org/10.1080/10494820.2022.2091606

Liu, C.-C., Cheng, Y.-B., & Huang, C.-W. (2011). The effect of simulation games on the learning of computational problem solving. *Computers & Education*, *57*(3), 1907–1918. https://doi.org/10.1016/j.compedu.2011.04.002

Liu, S. W., Chan, H. Y., & Hou, H. T. (2021). The development and evaluation of an online educational game integrated with real person-NPC mechanism for history learning. In *Proceedings of the 29th International Conference on Computers in Education. Asia-Pacific Society for Computers in Education*.

Liu, Y., Liu, D., Yuan, Y., & Archer, N. (2018). Examining situational continuous mobile game play behavior from the perspectives of diversion and flow experience. *Information Technology & People*, *31*(4), 948–965. https://doi.org/10.1108/ITP-02-2016-0042

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, *13*(4), 410. https://doi.org/10.3390/educsci13040410

Lukpat, A. (2023). CHATGPT banned in new York City public schools over concerns about cheating, and learning development. *Wall Street Journal*.

Makar, K., Bakker, A., & Ben-Zvi, D. (2015). Scaffolding norms of argumentation-based inquiry in a primary mathematics classroom. *ZDM*, *47*(7), 1107–1120. https://doi.org/10.1007/s11858-015-0732-1

Mhlanga, D. (2023). Open AI in education, the responsible and ethical use of ChatGPT towards lifelong learning. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4354422

Miller, L. M., Chang, C.-I., Wang, S., Beier, M. E., & Klisch, Y. (2011). Learning and motivational impacts of a multimedia science game. *Computers & Education*, *57*(1), 1425–1433. https://doi.org/10.1016/j.compedu.2011.01.016

Mogali, S. R. (2023). Initial impressions of ChatGPT for anatomy education. *Anatomical Sciences Education*, *17*(2), 444–447. https://doi.org/10.1002/ase.2261

Ng, C., Chen, Y., Wu, C., & Chang, T. (2022). Evaluation of math anxiety and its remediation through a digital training program in mathematics for first and second graders. *Brain and Behavior*, *12*(5), e2557. https://doi.org/10.1002/brb3.2557

Pallavicini, F., Pepe, A., & Mantovani, F. (2021). Commercial off-the-shelves video games for reducing stress and anxiety: A PRISMA systematic review (Preprint). *JMIR Mental Health*, *8*(8), e28150. https://doi.org/10.2196/28150

Pavlik, J. V. (2023). Collaborating with ChatGPT: Considering the implications of generative artificial intelligence for journalism and media education. *Journalism & Mass Communication Educator*, *78*(1), 84–93. https://doi.org/10.1177/10776958221149577

Pears, M., Yiasemidou, M., Ismail, M. A., Veneziano, D., & Biyani, C. S. (2020). Role of immersive technologies in healthcare education during the COVID-19 epidemic. *Scottish Medical Journal*, *65*(4), 112–119. https://doi.org/10.1177/0036933020956317

Perttula, A., Kiili, K., Lindstedt, A., & Tuomi, P. (2017). Flow experience in game based learning – A systematic literature review. *International Journal of Serious Games*, *4*(1), 57–72. https://doi.org/10.17083/ijsg.v4i1.151

Prensky, M. (2003). Digital game-based learning. *Computers in Entertainment*, *1*(1), 21–21. https://doi.org/10.1145/950566.950596

Rahman, M. M., & Watanobe, Y. (2023). ChatGPT for education and research: Opportunities, threats, and strategies. *Applied Sciences*, *13*(9), 5783. https://doi.org/10.3390/app13095783

Reed, J. M., & Ferdig, R. E. (2021). Gaming and anxiety in the nursing simulation lab: A pilot study of an escape room. *Journal of Professional Nursing*, *37*(2), 298–305. https://doi.org/10.1016/j.profnurs.2021.01.006

Ros, S., Gonzalez, S., Robles, A., Tobarra, L. L., Caminero, A., & Cano, J. (2020). Analyzing students' self-perception of success and learning effectiveness using gamification in an online cybersecurity course. *IEEE Access*, *8*, 97718–97728. https://doi.org/10.1109/ACCESS.2020.2996361

Rosenshine, B., & Meister, C. (1992). The use of scaffolds for teaching higher-level cognitive strategies. *Educational Leadership*, *49*(7), 26–33.

Sallam, M. (2023). The utility of ChatGPT as an example of large language models in healthcare education, research and practice: Systematic review on the future perspectives and potential limitations. *medRxiv*. https://doi.org/10.1101/2023.02.19.23286155

Sánchez-Ruiz, L. M., Moll-López, S., Nuñez-Pérez, A., Moraño-Fernández, J. A., & Vega-Fleitas, E. (2023). ChatGPT challenges blended learning methodologies in engineering education: A case study in mathematics. *Applied Sciences*, *13*(10), 6039. https://doi.org/10.3390/app13106039

Sobhanzadeh, M., Kalman, C. S., & Thompson, R. I. (2017). Labatorials in introductory physics courses. *European Journal of Physics*, *38*(6), 065702. https://doi.org/10.1088/1361-6404/aa8757

Spieler, B., Pfaff, N., & Slany, W. (2020). Reducing cognitive load through the worked example effect within a serious game environment. 2020 6th International Conference of the Immersive Learning Research Network (ILRN). https://doi.org/10.23919/iLRN47897.2020.9155187

Stevens, R. (2015). Role-play and student engagement: Reflections from the classroom. *Teaching in Higher Education*, *20*(5), 481–492. https://doi.org/10.1080/13562517.2015.1020778

Stokel-Walker, C. (2022). AI bot ChatGPT writes smart essays—Should academics worry? *Nature*. https://doi.org/10.1038/d41586-022-04397-7

Terwiesch, C. (2023). *Would ChatGPT get a Wharton MBA? A prediction based on its performance in the operations management course*. Mack Institute for Innovation Management at the Wharton School, University of Pennsylvania.

Urmeneta, C. E., & Walsh, S. (2017). Classroom interactional competence in content and language integrated learning. In A. Llinares & T. Morton (Eds.), *Applied linguistics perspectives on CLIL* (pp. 189–206). John Benjamins. https://doi.org/10.1075/lllt.47.11esc

Wardat, Y., Tashtoush, M. A., AlAli, R., & Jarrah, A. M. (2023). ChatGPT: A revolutionary tool for teaching and learning mathematics. *Eurasia Journal of Mathematics, Science and Technology Education*, *19*(7), em2286. https://doi.org/10.29333/ejmste/13272

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *17*(2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

Wouters, P., & van Oostendorp, H. (2016). Overview of instructional techniques to facilitate learning and motivation of serious games. In *Instructional techniques to facilitate learning and motivation of serious games* (pp. 1–16). Springer. https://doi.org/10.1007/978-3-319-39298-1_1

Xun, G., & Land, S. M. (2004). A conceptual framework for scaffolding III-structured problem-solving processes using question prompts and peer interactions. *Educational Technology Research and Development*, *52*(2), 5–22. https://doi.org/10.1007/BF02504836

Yang, Q.-F., Chang, S.-C., Hwang, G.-J., & Zou, D. (2020). Balancing cognitive complexity and gaming level: Effects of a cognitive complexity-based competition game on EFL students' English vocabulary learning performance, anxiety and behaviors. *Computers & Education*, *148*, 103808. https://doi.org/10.1016/j.compedu.2020.103808

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*(1), 39. https://doi.org/10.1186/s41239-019-0171-0

Zhai, X. (2022). ChatGPT user experience: Implications for education. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4312418

Zuo, T., Jiang, J., Spek, E., V. der Birk, M., & Hu, J. (2022). Situating learning in AR fantasy, design considerations for AR game-based learning for children. *Electronics*, *11*(15), 2331. https://doi.org/10.3390/electronics11152331