RESEARCH ARTICLE   **Open Access**

Developing a gamified artificial intelligence educational robot to promote learning effectiveness and behavior in laboratory safety courses for undergraduate students

Qi-Fan Yang[1†] , Li-Wen Lian[1†] and Jia-Hua Zhao[2*]

†Qi-Fan Yang and Li-Wen Lian
contributed equally to this work
and should be considered co-
first authors

*Correspondence:
jiahuazhaopum@gmail.com

[1] College of Education, Fujian
Normal University Cangshan
Campus, No. 8 Shangsan Road,
Cangshan District, Fuzhou,
Fujian, China
[2] Institute for Innovative
Learning, Mahidol University,
Nakhon Pathom, Thailand

Abstract

According to previous studies, traditional laboratory safety courses are delivered in a classroom setting where the instructor teaches and the students listen and read the course materials passively. The course content is also uninspiring and dull. Additionally, the teaching period is spread out, which adds to the instructor's workload. As a result, students become less motivated to learn. In contrast, artificially intelligent educational robots (AIERs), help students learn while lessening the workload on instructors by enhancing teaching strategies, using robots to substitute for teachers, giving students access to a variety of instructional content, and improving interaction with students through the use of intelligent voice interactions and Q&A systems to promote student engagement in learning. If the robot is used for a long time for learning, it may lead to a decrease in students' interest in learning. Therefore, this study introduces the GAFCC model (the theory-driven gamification goal, access, feedback, challenge, collaboration design model) as an instructional design model to guide the development of a gamified AIER system, aiming to improve students' motivation and learning effectiveness for laboratory safety courses. To test the effectiveness of the system, this study conducted an experimental study at a university in China in the summer of 2022. 53 participants participated in the research, with a random sample taken from each group. Each participant was able to choose the time of their free time to engage in the experiment. There were 18, 19, and 16 participants in experimental group 1, experimental group 2, and the traditional group, respectively. Students in experimental group 1 learned using the gamified AIER system, students in experimental group 2 learned on a general anthropomorphic robot system and the control group received traditional classroom learning. The experimental results showed that compared to the other two groups, the gamified AIER system guided by the GAFCC model significantly improved students' learning achievement and enhanced their learning motivation, flow experience, and problem-solving tendency. In addition, students who adopted this approach exhibited more positive behaviors and reduced cognitive load in the learning process.

Keywords: Artificial intelligence education robot, Gamification, GAFCC model, Laboratory safety, Higher Education

Introduction

Artificial intelligence (AI) development in recent years has opened up a wealth of opportunities for education (Zhang & Aslan, 2021). As research has shown, AI research includes the development of machines with a certain level of intelligence (Chen et al., 2020a), with robots attracting considerable attention in the field of education. Numerous studies have investigated the various roles that Artificial Intelligence Education Robots (AIERs) can play in teaching and learning, including those of teaching assistants (Smutny & Schreiberova, 2020), support tools for teachers (Reyes et al., 2021), and tools for a personal consultation (Muniasamy & Alasiry, 2020). Integrating learning content into AIERs can reduce students' isolation while learning, improve communication skills, and increase their interest in learning (Chin et al., 2011; Lewis Johnson, 2001; Shorey et al., 2019; Wu et al., 2020). Meanwhile, all of these advantages are contingent on how well various types of AIERs are designed and implemented. Virtual and physical robots can be classified based on their forms (Pei & Nie, 2018), with virtual robots typically serving as educational tools for programming and platforms for innovative practices (Kelleher et al., 2007), and physical robots typically serving as supplementary teaching tools, smart teachers, and learning companions, among other things (Kasap & Magnenat-Thalmann, 2012; Tanaka & Matsuzoe, 2012). In particular, physical robots are able to reinforce learner learning behaviors, affective outcomes, and increase learning benefits when interacting with learners (Belpaeme et al., 2018), enhance higher-level interaction experiences (Pei & Nie, 2018), and produce measurable cognitive learning outcomes (Leyzberg et al., 2012). However, when applied to education, AIERs have remained mainly in STEAM education, language education, and special education (Pei & Nie, 2018; Scaradozzi et al., 2019), with few studies integrating it into laboratory safety education.

The laboratory environment is an important part of university teaching and scientific research in which poor safety education is a global phenomenon (Ayi & Hon, 2018). University laboratory accidents have increased over the past 20 years (Bai et al., 2022b), student interest in laboratory safety education is low, and traditional teaching methods have diluted students' attitudes toward laboratory safety issues (Ménard & Trant, 2020). In addition, laboratory safety education is also very restricted by time and location, and any special circumstances encountered may easily increase the workload of the instructor. However, some previous research has also addressed the limitations of the current AIERs, for instance, showing that students' interest in using robots for prolonged practice decreases (Fryer et al., 2017). Additionally, guided strategies are frequently used in studies involving educational robots, while none of the others have any chosen strategies (Hwang & Chang, 2021).

Among the various learning strategies, students have a positive attitude toward game-based learning in courses or in informal learning (Wallace et al., 2010). Some researchers define gamification as increasing student engagement in learning by incorporating gamified elements into non-game educational settings (Hanus & Fox, 2015). Gamification is recognized as the most effective way to promote student learning and increase student learning enjoyment (Ge, 2018; Hew et al., 2016). Achievement unlocking in gamification elements is a successful method that can keep users

motivated over time (Groening & Binnewies, 2019). Furthermore, gamified AI is becoming an increasingly important area (Yannakakis & Togelius, 2018).

Past researchers have incorporated gamification strategies in the teaching of multiple disciplines, such as mathematics (Hung et al., 2014), language (Chu et al., 2022a), and science (Sung & Hwang, 2013), and findings have demonstrated the effectiveness of gamification strategies in subjects. However, few studies have applied gamification strategies to the development of AIERs in the field of laboratory safety. To increase students' learning motivation in laboratory safety courses, this study combines AIER with gamified elements to develop gamified laboratory safety courses that allow learners to interact with the robot when it is presented in a game-like manner. Therefore, a gamified AIER instructional system was proposed for college students to learn laboratory safety. To evaluate the validity of the approach, an experimental study was conducted at a university to answer the following research questions:

(1) Can students using the gamified AIER systems have better learning achievement than those who use the general anthropomorphic robot systems and traditional instruction?
(2) Can students using the gamified AIER system have a lower cognitive load than those who use the general anthropomorphic robot system and traditional instruction?
(3) Can students using the gamified AIER system show higher learning motivation than those who use the general anthropomorphic robot system and traditional instruction?
(4) Can students using the gamified AIER system have a better flow experience than those who use the general anthropomorphic robot system and traditional instruction?
(5) Can students using the gamified AIER system demonstrate better problem-solving tendency than those who use the general anthropomorphic robot system and traditional instruction?
(6) What are the differences in learning behaviors between students using the gamified AIER instructional system and students using the general anthropomorphic robot system?

Literature review

Artificial intelligence education robot

Robots are an innovative learning tool (b; Chu et al., 2022a,) that can modify education and support student learning in different learning contexts (Evripidou et al., 2020). Many schools also introduce educational robots as an innovative learning environment that fosters higher-level thinking skills such as situational awareness and critical thinking (Blanchard et al., 2010). Robots in educational settings can potentially support the development of students' skills such as problem-solving and collaboration skills (Chevalier et al., 2020). Educational robots are a powerful, flexible teaching and learning tool (Alimisis et al., 2009). It can be used to help students learn in a variety of subjects, including language, math, engineering, and STEAM (Lin et al., 2022; Sophokleous et al.,

2021; Sullivan & Bers, 2018; Zhong & Xia, 2020); in addition, it can improve students' social, programming, and computational thinking abilities (Bers et al., 2014; Diago et al., 2022). Meanwhile, using robotics in the classroom has huge potential to enhance classroom instruction (Benitti, 2012; Chandra et al., 2020; Papert, 1980).

With the development of AI, the application of AI in education has received much attention from researchers (b; Chu et al., 2022a) and is considered to have an important role in education (Limna et al., 2022). Researchers have noted that the integration of AI technologies has the potential to improve the learning efficiency of learners (b, c; Chen et al., 2020a; Hwang et al., 2020). For example, Timms (2016) claimed that anthropomorphic robots could communicate with humans through visual, auditory, and sensory systems to facilitate human–machine interaction; Yang and Zhang (2019) noted that incorporating AI technologies into robots could allow them to take on the role of teachers primarily. Additionally, artificial intelligence has significantly advanced technology for robots, such as speech recognition systems (Benkerzaz et al., 2019) and picture recognition systems (Sun et al., 2022). In the past, the majority of research on educational robotics has focused on the education of children (Fridin, 2014; Kewalramani et al., 2021). The use of teaching robots in higher education is currently generating significant interest from scholars (Maximova & Kim, 2016). All of the aforementioned studies have shown how effective AI robots are in the classroom, but Chu et al., (2022a, b), claimed that when it comes to learning strategies, AIERs most frequently use hybrid and problem-solving-related strategies. Fewer studies have used game-based strategies to encourage students' participation in AI robot learning activities. Therefore, it is necessary to develop game-based or gamified AIERs.

Gamification

Gamification is the use of tools and mechanisms, aesthetics, and game thinking to make people more engaged and motivated to do specific behaviors (Kapp, 2012). Deterding et al. (2011) defined "gamification" as the use of game elements in a non-game environment and noted that it has been used in educational settings. In recent years, researchers have used gamification methods in various educational applications (de-Marcos et al., 2016; Domínguez et al., 2013; Zhao et al., 2021). And the most notable advantages of using gamification for learning were the growth in students' attitudes, engagement, and performance. Points, badges, leaderboards, levels, feedback, and images were identified as important game elements for use in higher education (Subhash & Cudney, 2018). In a university programming course, Kasahara et al. (2019) used gamified elements and discovered that this method encouraged students to independently generate high-quality code and that students showed a higher willingness to learn. Hsu and Wang (2018) found that gamification helped improve the algorithmic thinking skills of elementary school students and enhanced their engagement experience and willingness to participate.

Nevertheless, some drawbacks of gamification have been identified in previous studies, including the lack of theoretical explanations to describe the link between gamification and motivational effects (Sailer et al., 2017; Seaborn & Fels, 2015), insufficient details given about the process and context of the gamified application (Falkner & Falkner, 2014; Hamari et al., 2014), and insufficient evidence left for the effectiveness of gamification (Hamari, 2017). Huang and Hew (2018) proposed a theory-driven gamification

Yang *et al. Int J Educ Technol High Educ*   (2023) 20:18

Page 5 of 31

model: the GAFCC model, which has been proven by researchers' experimental studies to address the above gamification shortcomings. For instance, Huang et al. (2019) employed the GAFCC model to explore the effects of gamification on students' online interaction patterns and peer feedback, and the results of the study showed positive feedback, which also provided empirical evidence. Also, Huang and Hew (2021) tested the validity of the GAFCC model again and refined it, concluding through a three-year experimental study that learning from gamified courses using the GAFCC model was satisfied with the overall learning design and that gamified courses promote student learning.

Past studies have pointed out that the most commonly used theories in gamification research are self-determination theory and flow theory (Kalogiannakis et al., 2021; Osatuyi et al., 2018). These two theories have been widely used in gamification research in educational settings (Zainuddin et al., 2020). Nadi-Ravandi and Batooli (2022) concluded through meta-analysis and systematic evaluation that social, cognitive, and behavioral theories dominate the theoretical frameworks used in game development. Thus, the GAFCC gamification design model uses self-determination theory (Ryan & Deci, 2000), flow theory (Csikszentmihalyi, 1978), goal setting theory (Locke & Latham, 2002), social comparison theory (Festinger, 1954), and behavioral reinforcement theory (Skinner, 1953) as the basic theoretical support for feasibility (Huang & Hew, 2018).

The gamified AIER system

Structure of the gamified AIER systems

The gamified AIER system, as displayed in Fig. 1, was created using the GAFCC model and consisted of four modules (a learning content module, an interactive practice module, a gamified learning module, and a learning material display module), as well as four databases (a speech recognition database, an image recognition database, a learning
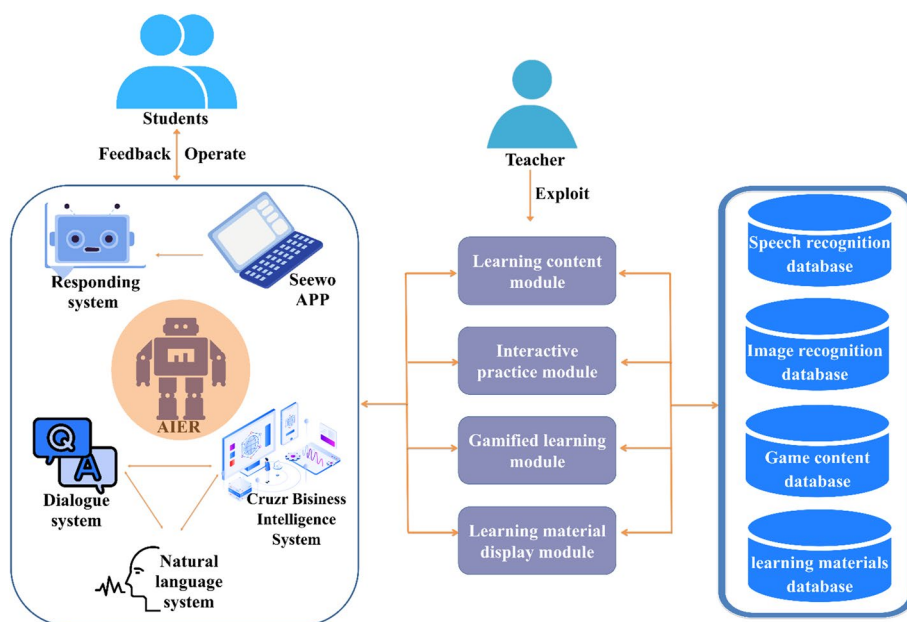

Fig. 1 System structure drawing

material database, and a game content database). When students interact with robots, the dialogue system and responding system are provided by the Cruzr Business Intelligence System and Seewo APP, respectively. By building the models and databases in the system, the teacher, on the other hand, provides the students with access to learning resources. The learning content module is for students to learn content to assist them in acquiring knowledge before entering the game. The learning material display module is to give students different forms of knowledge presentation modes, such as video, voice interaction, drawings, and text. The interactive practice module provides learning tasks for students. Following completion of the learning content module, corresponding questions are presented for students to answer either cooperatively or competitively in the interactive practice module. According to the progression of learning, the gamified learning module calculates scores throughout the learning process, selects the winner, and grants students the necessary badges. Additionally, four databases with various functions support the four modules and assist in storing and managing the data, improving the system's usability for students.

This AIER's intelligent dialogue system includes the functions of voice recognition and interaction, face recognition, face following, sound source location steering, etc. The AIER camera could automatically switch from standby to wake-up mode when it spots a person's face and will then take the initiative to do face recognition, broadcast greetings, and launch the home page. Fire safety, electrical safety, information security, and regulations are the four question categories that make up the learning content in the customized Q&A interface. Each question is structured with 4–6 fuzzy phrases, and each question category comprises 15–20 precise questions and answers. This allows students to ask inquiries with ease, using vocabulary they are familiar with, as shown in Fig. 2.

It is worth noting that in this system, all the course learning content has been organized into the aforementioned categories with associated questions and answers, creating a database. Students can, on the one hand, ask the robot questions based on the data displayed in the system interface or if they run into problems. The AIER can quickly recognize the user's language when they communicate with the robot, identify vague or precise keywords through a natural language processing system, and search through a database of knowledge materials created by the system developer, displaying corresponding videos, pictures, texts, voices, and body movements as feedback. On the other hand, if students asked questions that were outside the scope of the existing database, AIER would automatically network into an online search and provide the best response. In short, students will receive accurate replies when the keywords in their questions match or are relatively close to the content of the database, or conversely, the AIER system will search for the answers to provide intelligent solutions.

Content and framework of the gamified AIER system

As shown in Fig. 3, Huang and Hew (2018) summarized motivation needs into five fundamental components: goals, access, feedback, challenge, and collaboration. They also provided a brief explanation of the motivation theory and key components that support motivational experiences in the GAFCC model. These five fundamental components can be expanded upon using gamified aspects like points, breakout, and
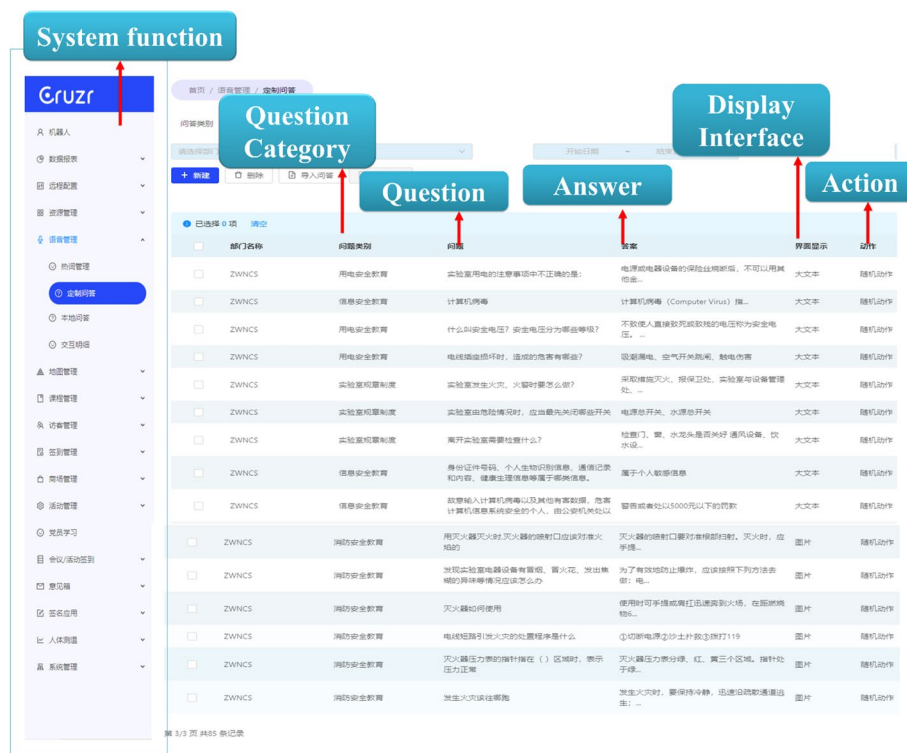
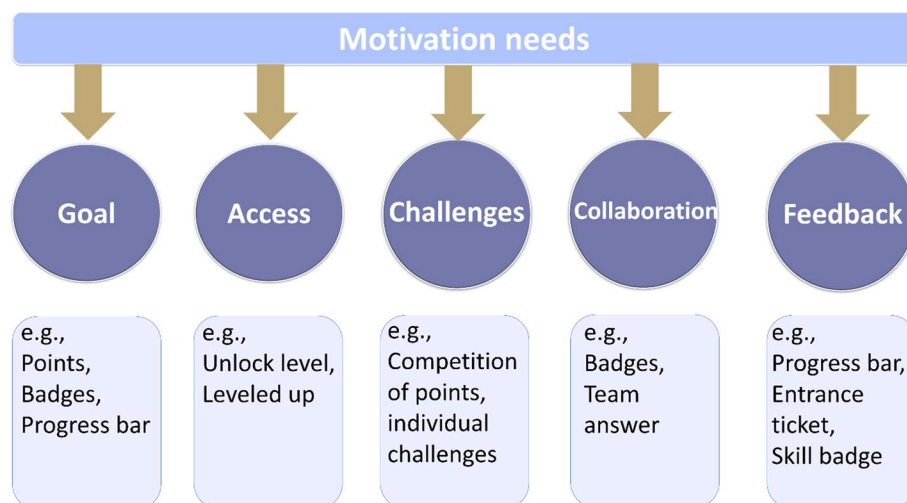**Fig. 2** AIER interactive system development interface

Fig. 3 Gamified design of the GAFCC model (adapted from Huang & Hew, 2018)

competitiveness. For instance, the goal element is primarily presented as points and skill badges; the access element can keep boosting students' motivation by unlocking levels; the challenge and collaboration elements are primarily presented as points and competition; the collaboration is presented as cooperative answers and badges; and the feedback element is presented as admission tickets and skill badges, etc.

Yang *et al. Int J Educ Technol High Educ* (2023) 20:18

Page 8 of 31

Design process for the gamified AIER system

The development systems for the application of gamified AIER on laboratory safety courses are mainly Cruzr Business Intelligence System and Seewo APP, with a design procedure following five phases (Huang & Hew, 2018; Huang et al., 2019):

Step 1: Examining the learning objectives, learner background, and technical competencies.

The primary learning goals of the laboratory safety course are to increase students' awareness of safety when performing laboratory operations and to advance their understanding of rules, laboratory practices, and safety precautions. Previous research has revealed that learners are uninterested in lectures on laboratory safety and do not view the material as being essential to their profession (Ménard & Trant, 2020). Moreover, the course is highly repetitive and represents a high workload for the instructor. Therefore, there is a need to identify an effective way to encourage students to take laboratory safety courses. With the development of AI, gamified plugins and robots could be a new approach to solve this problem, increasing student engagement, reducing instructors' teaching load, and improving the quality of teaching.

Step 2: Determining the motivating factors

Gamified elements are regarded as learning objectives in this study. Students could unlock levels by answering questions while they learn, with system recognition and timely feedback. Using challenges to encourage higher engagement and collaborative learning enhances the connection between learners.

Step 3: Matching motivators, game mechanics, and learning activities, as in Figure 4.
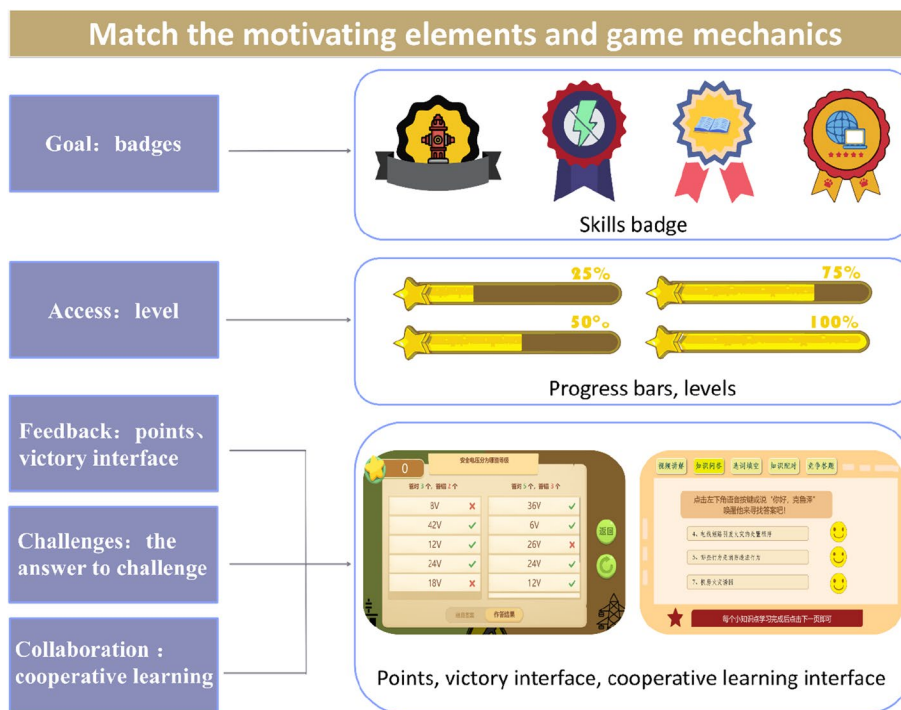

Fig. 4 Match the motivation needs and game mechanics

- Goal

Once learners have clarified their long-term and short-term goals, the system can use gamified elements such as badges to encourage students to set appropriate goals (Huang & Hew, 2018). The long-term objective of this study is lab entrance, while the short-term objective is to earn various badges by responding to questions and making breakthroughs. As a result, there are four different badges in this lesson: the rules and regulations badge, the internet safety badge, the electricity safety badge, and the badge for fire safety. The robot will dance and celebrate for students when they have successfully gathered all the badges and have gained admission to the lab.

- Access

The gamified AIER offers learners appropriate challenges, and students are led from easy to challenging by unlocking levels. In the gamified AIER, which leverages storyline-driven (i.e., obtaining access to labs) to boost student engagement, Fig. 5 depicts the beginning of a level. The system is set up with a variety of learning resources, including video lectures, tests, word searches, knowledge matching, and competitive quizzes. The system interface displays 0 points prior to the start of a level, and each time the learner completes a level, he or she can view the current progress and receive a skill badge.

Fig. 5 System display interface

- Feedback

   Praise and quick feedback might help students feel more motivated and driven to continue trying. Ongoing feedback helps students gain insight into their goals and follow their success (Hassan et al., 2021). In addition to rewarding students for learning at different levels, badges for different skills provide feedback on learner effort and skill acquisition. Students also get immediate feedback for each right or wrong response in the form of points, audible effects, and symbol changes on the answer screen. The system plays sound effects, presents a reward screen, and informs which student has won the challenge all at once. Both students can view the correct or incorrect responses to their questions for reflection and knowledge consolidation. Additionally, voice interaction buttons are always accessible and movable within the UI of the gamified AIER. The robot can be called by the student by saying "Hello, Cruzr," allowing them to receive timely feedback without interfering with the learning interface.

- Challenge

   Among the gamified elements, challenges are an essential part. Challenges can provide a platform for users to demonstrate their competence and success (Kyewski & Krämer, 2018). As they study, students are expected to provide answers to those questions. On the left and right sides of the same screen, two students participate in the competitive answer screen. The system's points are modified in real time so that it will give fast feedback on students' performance.

- Collaboration

   Bai et al. (2022a; b) stated that collaborative tasks should allow learners to communicate with each other more frequently. Interaction among students can help them strengthen their connection and communication with others. In the gamified AIER, the word selection interface and the knowledge matching interface require students to answer questions cooperatively, and students can complete these two parts of the content by discussing with each other or continuing to ask questions to the robot to promote cooperative learning among students (Fig. 6).

Step 4: Launching design

After matching motivational elements, game mechanics, and learning activities, the gamified design would be implemented.

Step 5: Evaluating the design

As the gamified design is implemented, quantitative and qualitative data analysis would be conducted to evaluate the results.

Experiment design

Participants

There were 53 students from two classes at a Chinese university, with an average age of 20. The students from the two classes were sorted into three groups according to a randomization principle for the experiment. 18 students in Class A were in the GAIER (gamified artificial intelligence education robot) group, which used gamified strategies to learn on AIER; 19 students in Class B were in the AR (anthropomorphic robot) group, which did not use gamified strategies for learning with anthropomorphic robots; and 16

Fig. 6 Competitive and cooperative game interface

students in Class C were in the TI (traditional instruction) group, which used PPT and other learning materials for learning. And all three groups were instructed by the same laboratory safety instructor.

Instruments

The instruments for this study included student pre-and post-tests, pre-and post-questionnaires on cognitive load, learning motivation, mind-flow experiences, and problem-solving tendencies.

The purpose of the pre-test was to assess the students' prior knowledge, i.e., their learning of basic laboratory safety. It consisted of 14 single and multiple choice questions (42%), 5 judgment questions (10%), 4 fill-in-the-blank questions (12%), and two short answer questions (36%) out of 100 points. The purpose of the post-test is to assess student performance in learning activities with the same types of questions and scores as the pre-test. The tests were developed from a database of test questions from school laboratory safety education exams and were evaluated by two teachers with more than 10 years of experience in teaching laboratory safety to ensure that the pre and post-tests of the selected knowledge content adequately assessed students' learning performance. Both tests were of equal difficulty and scored a total of 100 points. Among the 23 questions that could be fully judged as correct or incorrect (e.g., multiple choice, fill-in-the-blank, and judgment questions), they were divided into three different scales based on the difficulty index (D), which were 0%-40% for difficult questions (6/23), 40%-70% for moderate questions (9/23), and 70%-100% for easy questions (8/23). Moreover, the total

discrimination index (d) for all questions ranged from 0.00 to 0.84, which was mainly focused in the interval of 0.33–0.52 (Salkind, 2017). The Cronbach's alpha values for the tests were 0.80 (Cortina, 1993).

The cognitive load questionnaire was adapted from Hwang et al. (2013) to understand the impact of the approach on students. The questionnaire was completed at the end of the learning activity. This questionnaire consists of 8 items with a Likert scale score of 7 (7 = strongly disagree; 1 = strongly agree). Five of the items were on mental load, and the other three were on mental effort. The Cronbach's alpha values for the two dimensions were 0.86 and 0.85, respectively, e.g., "The learning content in this activity is difficult for me," and "The teaching style or presentation of the material in this learning activity is difficult for me.

The learning motivation questionnaire was adapted from Wang and Chen (2010) based on Pintrich et al. (1991). This questionnaire consists of six items with a Likert scale of 5 (1 = strongly disagree; 5 = strongly agree), and the measurement dimensions are divided into intrinsic and extrinsic motivation, with Cronbach's alpha values of 0.79. This questionnaire was completed before and after the learning activity. Examples of questionnaire were "In this course, I prefer challenging material because I can learn new things" and "In this course, I prefer challenging material because I can learn new things". " and "Getting a good grade in this course is the most satisfying thing for me".

The flow experience questionnaire uses a scale developed by Pearce et al. (2005). This questionnaire was completed after the learning activity. This questionnaire consists of eight 5-point Likert scales (1 = strongly disagree; 5 = strongly agree), such as: "I am strongly engaged in this activity " and "I find this activity enjoyable". And the Cronbach's alpha value was 0.82.

Problem-solving tendency questionnaire was measured using a scale developed by Lai and Hwang (2014). This questionnaire consists of six items on a 5-point Likert scale (1 = strongly disagree; 5 = strongly agree) with Cronbach's alpha values of 0.78. This questionnaire was completed before and after the learning activity. Examples of this test are "I believe I am capable of solving the problems I encounter problems" and "I believe I can solve problems on my own."

The interview was modified from a measure developed by Hwang et al. (2009). It consisted of seven items to assess students' attitudes toward this gamified AIER learning (e.g., "How is this gamified AIER different from previous courses you have taken? What are the overall advantages of this learning method? What kind of help did you get with this approach? Please give an example"). To understand the impact of gamified AIER as an intervention on students, we randomly invited 10 students each from experimental group 1 and experimental group 2, for a total of 20 students, to participate in interviews to demonstrate their thoughts and feelings in the corresponding learning environments.

Coding scheme for learning behaviors

This study divided learning behaviors into three categories: active learning behaviors, passive learning behaviors, and question-answering behaviors to investigate the association between learning behaviors and learning achievement during gamified learning. The high-definition cameras are used to record student classroom behavior data. And

the students' behavior coding system was created by combining the game characteristics of this study with the coding schemes developed by Sung and Hwang (2018), Zhang et al. (2020), and Hwang et al. (2018).

As shown in Table 1, the behaviors of students actively reading the learning materials, recounting and memorizing the lessons, interacting with classmates, and asking questions regarding unsolved difficulties are examples of how learning behaviors are coded. Students read the questions before responding to them either cooperatively or competitively, depending on the system settings. For instance, when students read the question,

Table 1 Coding scheme for the students' behaviors

Classification	Code	Definition	Example
Active learning behavior	L1	Reading learning materials	Students read the learning materials
	L2	Replacement of learning materials	Students actively change learning materials after reading the current learning materials
	L3	Retelling	Students repeat and follow along as they study the learning materials
	L4	Memorizing	Students repeatedly read, recite, or stay for more than 10 s after studying the materials
	L5	Discussion	Students share and discuss among themselves what is relevant to this study
	L6	Asking questions	Students ask questions to the robot
	L7	Obtaining learning materials	Learning materials would be given by the robot after the student asked the robot questions
	L8	Revisit/access learning materials	Students revisit material they have already learned or ask the robot a new question to obtain learning material
	L9	Check out the feedback	Students actively check the feedback given by the system after answering the questions, which includes the total number of questions answered, the number of normal/incorrect answers, the analysis of the questions, and the standard answers
Passive learning behavior	P1	Other behaviors	Two students chatted about non-study related topics, played with their cell phones, and looked around during their studies
	P2	Emotional up and down	Students show higher emotions, excitement, joy, eagerness to try, etc
Answer behavior	A1	Reading the questions	Students read the questions in the system
	A2	Answer questions cooperatively	Students answer questions together
	A3	Answer the questions competitively	Students can answer the questions through their respective interactive buttons, or through timed PKs
	A4	Re-answer	Students answer again after giving a wrong response
	A5	Answer correctly	After students complete their answers, the robot gives feedback that the answer is correct
	A6	Answer incorrectly	After students complete their answers, the robot gives feedback that the answer is incorrect

it would be recorded as A1; when students can receive feedback from the robot on answering the question correctly, it would be recorded as A5; if the student answers incorrectly, feedback on the incorrect answer is given, which would be recorded as A6. The learning behaviors are recorded in the chronological order in which they occur.

Experimental process

Figure 7 illustrates the experimental procedure. Before the learning activity started, all students were asked to take a pre-test and pre-questionnaire for a total of 30 min. In the second week, the instructor spent 20 min introducing the learning objectives and rules for robot use, after which students started the learning activities using different learning strategies. Experiment 1 group, the GAIER group, used the gamified AIER system as a learning tool in the laboratory safety classroom. This group of students all worked in groups of two to learn together. The AR group, as experiment 2 group, also worked in groups of two to learn through the AR system in the classroom. The robot in the AR group still has a voice interaction system and a Q&A system, and it shares the same database of learning resources as the GAIER group, which can still interact with students. Yet, the interaction process is not gamified, and incentives like verbal praise, sound effects, and badges are not used. Students are provided feedback on whether their responses are correct after responding to the questions, but there are no points calculated in the question–answer interface. The control group, as the TI group, was asked to use the PPT and learning materials for traditional learning. All groups spent
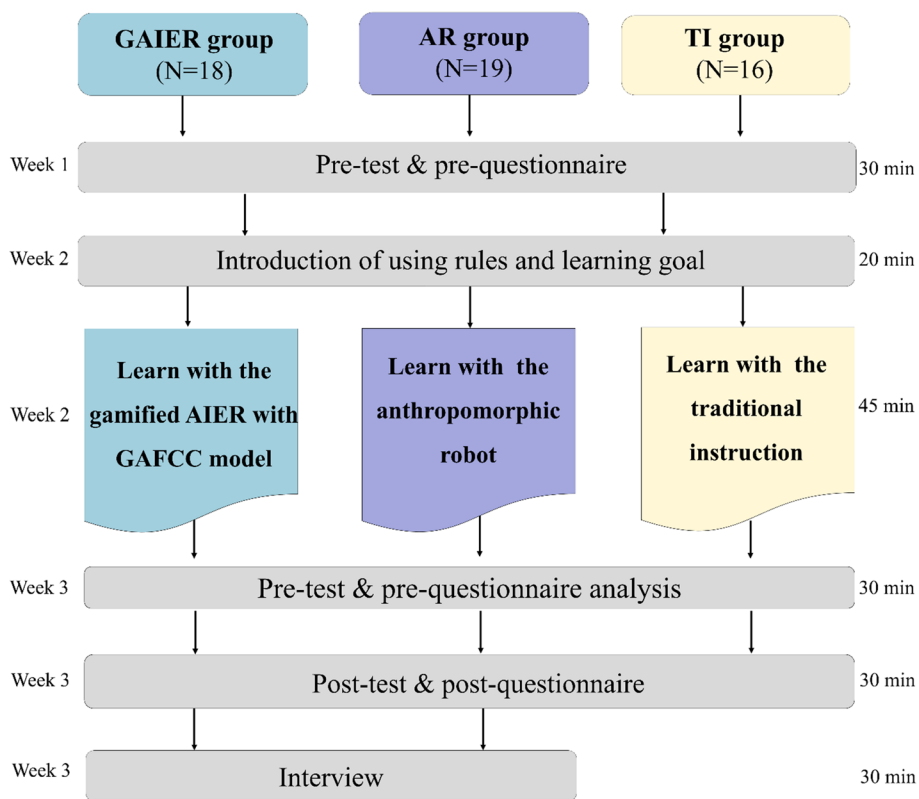

Fig. 7 The experimental procedure

45 min to complete the learning task. After the learning activity, all students were asked to take a post-test, which lasted 30 min. Finally, the researchers randomly selected 10 students from each class to participate in the interviews, and a total of 20 students were interviewed.

Experimental results

Learning achievement

To detect differences in learning achievement, pre-test scores were used as the covariate, groups as the independent variable, and post-test scores as the dependent variable. First, the Shapiro–Wilk test was used to test the normality of the data, and the result was 0.97, $p = 0.19 > 0.05$, indicating that the sample in this study was normally distributed. Also, it was determined that the Levene test for homogeneity of variance was not violated ($F = 2.94$, $p = 0.06 > 0.05$), indicating that the null hypothesis is valid and that the variances between the groups are equal. In addition, homogeneity of regression coefficients within groups was confirmed, $F_{(2,47)} = 2.45$, $p = 0.10 > 0.05$, indicating that analysis of covariance was appropriate. Therefore, a one-way ANCOVA analysis was performed.

Table 2 shows the results of the ANCOVA from the three groups of post-tests. The results showed that the differences in test scores between the three groups were significant($F_{(2,49)} = 5.928$, $p < 0.01$). In addition, post hoc analyses were conducted to examine specific differences in learning achievement between the experimental groups. The LSD (Least Significant Difference) test noted that comparing the adjusted mean of 83.57 for the GAIER group with the score of 69.25 for the TI group ($p < 0.01$), the GAIER group scored significantly higher than the TI group. Comparing the adjusted mean of 77.77 in the AR group and the score of 69.25 in the TI group ($p < 0.05$), the score of the AR group was significantly higher than that of the TI group. The results implied that students who learned with gamified AIER system and those who learned with anthropomorphic robot system had better learning outcomes than students who learned with traditional instruction.

cognitive load

A post-questionnaire was used to investigate the cognitive load of each group of students, with groups as the independent variable and post-test scores as the dependent variable, and ANOVA analysis was conducted.

The results of the ANOVA are provided in Table 3, with five items of mental load, $F = 12.07$, $p < 0.001$; three items of mental effort, $F = 8.41$, $p < 0.01$. Therefore, the overall scale showed there was a significant difference across the three groups

Table 2 Descriptive data and the ANCOVA result of the learning achievement for the three groups

Group	N	Pre-test Mean	Post-test				
			Mean	S.D	Adjusted mean	*F*	Post Hoc
(1) GAIER	18	55.28	86.06	8.21	83.57	5.928[**]	(1) > (3)
(2)AR	19	45.53	74.63	13.88	77.77		(2) > (3)
(3)TI	16	48.56	68.50	15.12	69.25		

[**]$p < 0.01$

Table 3 Descriptive data and the ANOVA result of the cognitive load for the three groups

Group	N	Mean	S.D	F	Post Hoc
(1)GAIER	18	6.15	0.12	13.028***	(1) > (3)
(2)AR	19	6.11	0.14		(2) > (3)
(3)TI	16	5.01	0.26		

*** $p < 0.001$

Table 4 Descriptive data and the ANCOVA result of the learning motivation for the three groups

Group	N	Pre-test Mean	Post-test				
			Mean	S.D	Adjusted mean	F	Post Hoc
(1)GAIER	18	3.31	3.96	0.44	4.06	3.864*	(1) > (3)
(2)AR	19	3.62	3.91	0.55	3.87		
(3)TI	16	3.65	3.67	0.54	3.61		

* $p < 0.05$

in the cognitive load ($F_{(2,50)} = 13.03$, $p < 0.001$). In addition, post hoc analyses were conducted to examine specific differences between the three groups, with the LSD test showing that the GAIER group scored significantly higher than the TI group, and the AR group also scored significantly higher than the TI group. The results noted that both students who learned using the gamified AIER system and those who learned using the anthropomorphic robot system expressed a lower cognitive load than those who learned using the traditional method.

Learning motivation

To detect learning motivation, pre-questionnaire on learning motivation was used as the covariate, the groups as the independent variable, and post-questionnaire on learning motivation as the dependent variable. First, the Shapiro–Wilk test was used to test the normality of the data, and the result was 0.99, $p = 0.81 > 0.05$, indicating that the sample in this study was normally distributed. Also, the post learning motivation questionnaire of the three groups of students was investigated, and the Levene's test to determine homogeneity of variance was not violated ($F = 0.87$, $p = 0.43 > 0.05$), indicating that the original hypothesis was valid and the variance between groups was equal. Meanwhile, homogeneity of regression coefficients was confirmed, hence, the use of analysis of covariance was appropriate ($F = 0.35$, $p = 0.71 > 0.05$).

As shown in Table 4, the ANCOVA results illustrated a significant difference in learning motivation among the three groups ($F_{(2,49)} = 3.86$, $p < 0.05$). In addition, post hoc analyses were conducted to examine specific differences in learning motivation between the three groups. According to the LSD test results, the GAIER group scored significantly higher than the TI group. The results demonstrated that

students who learned using the gamified AIER system were better motivated to learn than those who learned using traditional methods.

Flow experience

To investigate whether the combination of gamification elements and robotics could enhance students' flow experience, the experiment used a pre-questionnaire to investigate the flow experience of each group, with groups as the independent variable and post-questionnaire as the dependent variable, and ANOVA analysis was conducted. The results of the ANOVA are provided in Table 5.

The results showed that there was a significant difference between the three groups of cardiac flow experience ($F_{(2,50)} = 5.95$, $p < 0.01$). Also, post hoc analyses were conducted to examine specific differences in flow experience between the three groups. The LSD test revealed that the GAIER group scored significantly higher than the AR group, and the AR group also scored significantly higher than the TI group. The results indicated that students learning with gamified AIER system showed higher flow experiences than those learning with anthropomorphic robot system and those learning with traditional methods.

Problem-solving tendency

To test the students' problem-solving tendency, the pre-questionnaire on problem-solving tendency was used as the covariate, the groups as the independent variable, and the post-questionnaire on problem-solving tendency as the dependent variable. The Shapiro–Wilk test was used to examine the normality of the data, and the test result was 0.97, $p = 0.28 > 0.05$, indicating that the sample in this study was normally distributed. A post hoc questionnaire was used to investigate the problem-solving tendency of various students. The Levene's test to determine homogeneity of variance ($F = 0.47$, $p = 0.63 > 0.05$) was not violated, indicating that the original hypothesis was valid and the variance between groups was equal. In addition, homogeneity of regression coefficients was confirmed, indicating that the use of analysis of covariance was appropriate ($F = 2.72$, $p = 0.08 > 0.05$). Therefore, an ANOVA analysis was performed.

The results of the ANCOVA shown in Table 6 showed that there was a significant difference in problem solving tendencies between the three groups ($F_{(2,49)} = 5.29$, $p < 0.05$). Besides, post hoc analyses were conducted to examine specific differences in problem solving tendencies between the three groups. The LSD test revealed that the GAIER group scored significantly higher than the TI group, and the AR group scored significantly higher than the TI group. The results suggest that students who learned using

Table 5 Descriptive data and the ANOVA result of the flow experience for the three groups

Group	N	Mean	S.D	*F*	Post Hoc
(1)GAIER	18	4.34	0.11	5.948**	(1) > (2)
(2)AR	19	3.90	0.14		(1) > (3)
(3)TI	16	3.70	0.15		

**$p < 0.01$

Table 6 Descriptive data and the ANCOVA result of the problem-solving tendency for the three groups

Group	N	Pre-test Mean	Post-test				
			Mean	S.D	Adjusted mean	*F*	Post Hoc
(1)GAIER	18	3.93	4.31	0.32	4.28	5.293[**]	(1) > (3)
(2)AR	19	3.80	4.19	0.45	4.22		(2) > (3)
(3)TI	16	3.86	3.91	0.50	3.91		

[**]$p < 0.01$

the gamified AIER system and those who learned using anthropomorphic robot system had stronger problem-solving tendencies than those who learned using traditional instruction.

Analysis of learning behavior patterns

To explore the differences in learning behaviors between the two groups of students, behavioral sequence analysis was used. The z-values were calculated to evaluate the coded data of each group and to generate a table of adjusted residuals for the students' behavioral patterns. Table 7 shows the adjusted residual table for the GAIER group. A z-value greater than 1.96 indicates that the sequence is statistically significant (Bakeman & Gottman, 1997). GSEQ 5.1 developed by Quera et al. (2007) was used for sequence analysis in this study.

Figure 8 illustrates the behavioral patterns of the GAIER group; the number on each line represents the z value of the sequence, and the direction of each line denotes the transfer direction. Additionally, to separate the significance level of the sequence, the thicker line indicates that the z-value of the behavior was greater than 8.00, while the thinner line indicates that it was greater than 1.96 but less than 8.00. According to the figure, L6 → L7 indicates that the students questioned the robot to receive the learning materials; L7 → L1 indicates that the students read the learning materials after obtaining them; L1 → P1 indicates that students watched the learning materials while engaging in non-learning behaviors; L1 ↔ L2 indicates students independently changed the learning materials; L1 → L3 and L1 → L4 indicate students retold and memorized the learning materials; L3 → A1 and L4 → A1 indicate students began reading the questions as they entered the question–answer stage. Cooperative and competitive responses make up the majority of the answering stage. A1 → A3 demonstrates that after reading the questions, students proceed on to the competition stage; A3 → L9 demonstrates that following the competition, students actively checked the evaluations; A3 → P2 demonstrates that students had emotional ups and downs while competing; Students read the questions and then discussed them, as shown by A1 → L5; Students interacted before working together to answer the questions, as seen by L5 → A2; A2 → A5 indicates that students answered correctly in the process of cooperative answer; A2 → A6 indicates that students answered incorrectly in the process of cooperative answer; A6 ↔ A4 indicates that students would re-answer after answering incorrectly, but the possibility of answering incorrectly will still occur after re-answering; Students who re-answered correctly received an answer grade of

Table 7 Adjusted residual table of the GAIER group

| Z	L1	L2	L3	L4	A1	L5	A2	A3	A4	L8	A5	A6	L6	L7	P1	L9	P2
L1	0	5.52*	13.7*	7.31*	4.39*	0.48	− 3.89	− 3.67	− 1.92	− 0.92	− 3.97	− 2.07	0.85	− 8.96	2.81*	− 1.5	− 1.76
L2	5.62*	0	− 0.77	− 0.39	− 0.88	− 0.89	− 0.59	− 0.56	− 0.29	− 0.14	− 0.61	− 0.32	− 1.38	− 1.37	− 0.26	− 0.23	− 0.71
L3	− 4.91	− 0.77	0	− 1.24	3.11*	− 1.14	− 1.87	− 1.76	− 0.92	− 0.44	− 1.91	− 0.99	13.18*	− 4.3	0.53	− 0.72	− 1.71
L4	− 2.53	− 0.39	− 1.24	0	5.58*	0.1	− 0.96	− 0.91	− 0.48	− 0.23	− 0.98	− 0.51	4.36*	− 2.22	− 0.42	− 0.37	− 1.14
A1	− 5.66	− 0.88	− 2.77	− 1.43	0	16.29*	− 0.57	21.91*	− 1.07	− 0.51	− 2.2	− 1.15	− 4.72	− 4.96	0.25	− 0.83	− 2.56
L5	− 4.31	− 0.89	− 0.72	0.1	0.45	0	20.85*	− 2.06	− 1.08	5.88*	− 1.2	− 0.21	1.2	− 5.02	0.23	0.46	− 0.8
A2	− 3.81	− 0.59	− 1.87	− 0.96	− 2.15	− 2.18	0	− 0.58	0.75	− 0.34	27.24*	10.18*	− 3.35	− 3.34	− 0.63	− 0.56	− 1.08
A3	− 3.6	− 0.56	− 1.76	− 0.91	0.75	0.7	− 1.37	0	− 0.68	− 0.32	− 0.62	− 0.73	− 3.17	− 3.15	− 0.59	3.42*	20*
A4	− 1.89	− 0.29	− 0.92	− 0.48	− 1.07	− 1.08	− 0.72	− 0.68	0	− 0.17	10.8*	13.02*	− 1.66	− 1.65	− 0.31	− 0.28	− 0.85
L8	− 0.9	− 0.14	− 0.44	− 0.23	− 0.51	− 0.51	8.83*	− 0.32	− 0.17	0	− 0.35	− 0.18	− 0.79	− 0.79	− 0.15	− 0.13	− 0.41
A5	− 2.18	− 0.6	− 1.89	− 0.97	2.01*	− 2.2	− 1.47	− 1.38	− 0.73	− 0.35	0	− 0.78	3.25*	− 3.38	1.02	1.29	9.68*
A6	− 1.41	− 0.32	− 0.99	− 0.51	− 1.15	− 1.16	− 0.77	− 0.73	29.1*	− 0.18	0.55	0	− 1.1	− 1.78	− 0.33	− 0.3	0.25
L6	− 8.82	− 1.38	− 4.32	− 2.22	− 4.98	− 5.04	− 3.35	3.17	− 1.66	− 0.79	− 3.42	− 1.78	0	29.19*	− 1.45	− 1.3	− 3.67
L7	24.81*	− 1.37	− 4.3	− 2.22	− 4.96	− 5.02	− 3.34	− 3.15	− 1.65	− 0.79	− 3.41	− 1.78	− 7.72	0	− 1.45	− 1.29	− 3.97
P1	− 0.89	− 0.26	0.53	− 0.42	6.15*	− 0.94	− 0.63	1.17	− 0.31	− 0.15	− 0.64	2.72*	− 1.45	− 1.45	0	− 0.24	− 0.75
L9	− 0.63	− 0.23	− 0.72	− 0.37	0.49	1.77	− 0.56	− 0.53	− 0.28	− 0.13	− 0.57	− 0.3	1.49	− 1.29	− 0.24	0	0.93
P2	− 0.28	− 0.67	− 0.47	− 0.08	− 0.02	2.29*	− 1.62	− 1.53	0.53	− 0.38	− 0.99	− 0.86	3.99*	− 3.73	− 0.7	6.14*	0

*z > 1.96
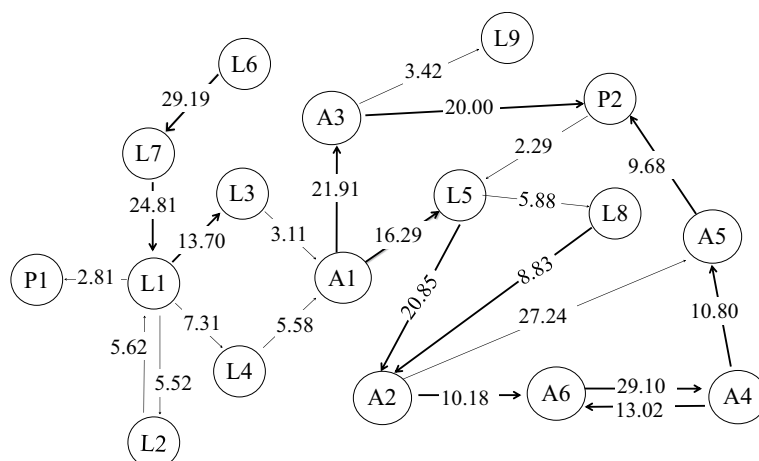
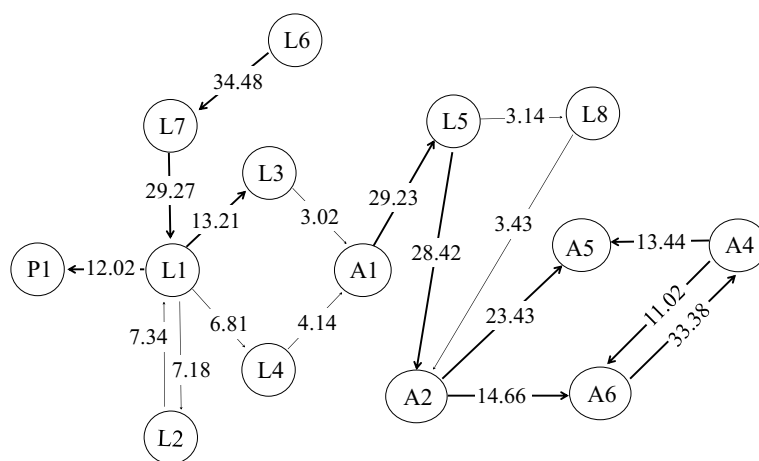**Fig. 8** Behavioral patterns of the GAIER group

Fig. 9 Behavioral patterns of the AR group

A4 → A5; A5 → P2 denotes that after providing the correct response, students' emotions changed; Following the conversation and discussion, students reread the learning materials as shown by the L5 → L8; L8 → A2 indicates that students continued to answer the questions after revisiting the material.

Similarly, this study also analyzed the behavior of the AR group, as shown in Fig. 9 and Table8. The AR group was basically similar to the experimental group in terms of learning behavior, which was L6 → L7 → L1, or L1 ↔ L2, i.e., autonomously asking questions to the robot to obtain material for learning or autonomously changing the learning material given in the system. The learners also had the behavior of memorizing and retelling the learning materials before answering during the learning process. In terms of answering behavior, there was only cooperative answering in the AR group, so the behavior of the vast majority of the groups was from A1 → L5, reading the questions and then exchanging and discussing them; then L5 → A2, exchanging and then cooperating to answer the questions. l5 → L8 was the behavior of relearning when students found that they could not answer after reading the questions.

Table 8 Adjusted residual table of the AR group

| Z	L1	L2	L3	L4	A1	L5	A2	A3	A4	L8	A5	A6	L6	L7	P1	L9	P2
L1	0	7.18*	13.21*	6.81*	1.6	− 1.4	− 5.82	0	− 3.29	− 1.24	− 5.85	− 3.58	4.49*	− 6.39	12.02*	0	6.68*
L2	7.34*	0	− 0.53	− 0.24	− 1.15	− 1.22	− 1.14	0	− 0.6	− 0.23	− 1.07	− 0.66	− 1.17	− 1.17	− 0.48	0	− 0.29
L3	− 2.8	− 0.53	0	− 0.5	3.02*	− 2.05	− 2.36	0	− 1.25	− 0.47	− 2.21	− 1.36	11.58*	− 2.42	− 1	0	− 0.6
L4	− 1.28	− 0.24	− 0.5	0	4.14*	− 0.16	− 1.08	0	− 0.57	− 0.21	− 1.01	− 0.62	1.98*	− 1.11	− 0.46	0	− 0.27
A1	− 6.12	− 1.15	− 2.37	− 1.09	0	29.23*	− 2.44	0	− 2.72	0.09	− 4.83	− 2.96	− 4.79	− 5.28	− 2.18	0	− 1.31
L5	− 6.5	− 1.22	− 1.59	− 1.16	− 3.62	0	28.42*	0	− 2.48	3.14*	− 5.13	− 3.15	− 1.45	− 5.61	− 1.81	0	− 1.39
A2	− 5.88	− 1.14	− 2.36	− 1.08	− 4.17	− 5.48	0	0	− 2.71	− 1.02	23.43*	14.66*	− 4.77	− 5.26	− 1.64	0	− 0.43
A3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A4	− 3.21	− 0.6	− 1.25	− 0.57	− 2.72	− 2.89	− 2.71	0	0	− 0.54	13.44*	11.02*	− 2.77	− 2.77	− 1.15	0	− 0.69
L8	− 0.24	− 0.23	− 0.47	− 0.21	− 1.02	− 0.03	3.43*	0	3.35*	0	− 0.96	− 0.59	− 1.04	− 1.04	− 0.43	0	− 0.26
A5	0.6	− 1.03	− 2.13	− 0.98	14.37*	− 4.95	− 4.64	0	− 2.45	− 0.92	0	− 2.66	5.02	− 4.75	− 1.39	0	− 0.23
A6	− 2.8	− 0.66	− 1.36	− 0.62	− 0.97	− 2.77	− 2.95	0	33.38*	4.8*	− 2.34	0	− 3.02	− 3.02	− 1.25	0	− 0.75
L6	− 6.24	− 1.17	− 2.42	− 1.11	− 5.28	− 5.61	− 5.26	0	− 2.77	− 1.04	− 4.92	− 3.02	0	34.48*	− 2.22	0	− 1.34
L7	29.27*	− 1.17	− 2.42	− 1.11	− 5.28	− 5.61	− 5.26	0	− 2.77	− 1.04	− 4.92	− 3.02	− 5.38	0	− 2.22	0	− 1.34
P1	− 2.58	− 0.48	− 1	− 0.46	7.9*	− 2.32	− 2.17	0	− 1.15	− 0.43	− 1.47	− 1.25	5.61	− 2.22	0	0	− 0.55
L9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P2	− 1.55	− 0.29	1.13	− 0.27	0.43	− 1.39	− 1.3	0	− 0.69	− 0.26	− 0.3	− 0.75	6.37	− 1.34	− 0.55	0	0

*z > 1.96
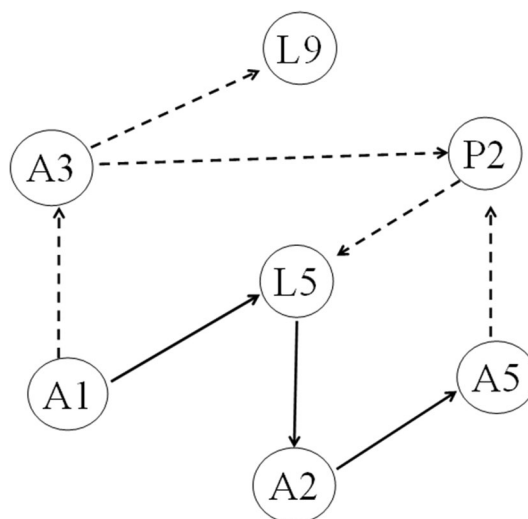
Fig. 10 Behavioral learning patterns of the students in the two groups

The results of the behavior patterns of the two groups were compared, as shown in Fig. 10, with the solid lines representing behaviors common to both the GAIER and AR groups and the dashed lines representing behaviors specific to the GAIER group. Several differences in learning behaviors were found between the two groups, with individually significant and notable behaviors occurring only in the GAIER group, e.g., only students in the GAIER group exhibited the behaviors of competition (A3), checking feedback (L9), and emotional ups and downs (P2), which may be key factors in the superior learning of Experimental Group 1 over Experimental Group 2. For example, the A3 → P2 behavior reflects the advantage of adding gamification elements to the GAIER group, i.e., elements such as competition, points, and sound effects of victory or defeat made students show significant emotional actions and language during learning; the significance of A3 → L9 indicates that students were more willing to check their mistakes and actively view feedback after competition; The significance of A5 → P2 behavior is due to the fact that the GAIER group has certain storyline and levels, and students can get badges after completing the study, which makes their learning motivation enhanced and more interesting. Also, when students' learning emotions gradually rise, it also promotes communication among students, which is the reason for the significant P2 → L5 behavior. During the system setup of the GAIER group, levels, badges, and task progress bars were added to help students grasp the overall learning progress and improve their attention. Moreover, the competitive answer activity included countdown, points, sound effects, and special effects, which brought students a certain sense of tension and better motivated their desire to win and lose. At the same time, the competition activity requires two students to participate in the learning at the same time, and students need to memorize the learning material carefully in order to win, which in turn enhances students' active learning.

Table 9 Coding items and number of occurrences of the interview results

Classification	Code	GAIER group	AR group
Learning achievement	Improve learning efficiency	2	2
	Make learning easier	1	2
	Motivate them to study hard	6	1
Learning motivation	Competition motivates learning to win	5	0
	Stimulate interest and engagement in learning	8	6
Cognitive load	Interact more and communicate easily	7	3
	Provide timely feedback	2	4
	Control progress and repeatable course content	3	3
Flow experience	Competitive and enjoyable	4	0
	It is very fun	6	4

Analysis of interview

The findings from the analysis of the interviews, including the codes and the frequency of occurrences, are presented in Table 9. After comparison, "better learning achievement" (N=9), "increased motivation" (N=13), "lower cognitive load" (N=12), and "enhanced flow experience" (N=10) were among the learning benefits obtained.

In terms of achievement improvement, S109 from the GAIER group said: "Studying alongside my classmates sparked interest and drive in studying, I became competitive because of the competition, and I made a concerted effort to memorize information and learn more than my opponents. I can't get distracted since the robot is less monotonous and more interactive". "I was more engaged in asking questions, recalling the information, and piquing my curiosity about learning", according to S206 from the AR group. In conclusion, students in both the GAIER and AR groups thought that the robot's role as a teacher helped them learn to boost their learning motivation and achievement. However, the difference is that students in the GAIER group are more serious about learning, while the learning mode in the AR group is not gamified, and the learning period is sufficiently long to feel tedious, as S210 in the AR group noted: "Although the robot provided prompt responses, it felt more lively than a traditional classroom. But when I was learning, I thought it was less exciting and more routine". Also, S205 said: "It was fun But it was repetitive and I had to click on the screen all the time. I hope new features can be added to make it more interesting".

In terms of enhanced learning motivation and flow experience, S105 in the GAIER group said: "GAIER made me interact with my peers competitively to arouse our interest. Having someone to compete with while answering questions would make learning more serious and clearer than remembering alone". "The competition sparked my curiosity and determination to learn", S202 in the AR group added. According to frequency statistics, students in the GAIER group showed greater motivation and interest in studying than those in the AR group.

In terms of cognitive load, S104 in the GAIER group described that: "GAIER was freer, I could manage the pace and speed of learning on my own, and I could play back anything I couldn't understand". S202 from the AR group stated that "the robot would direct me and answer my new inquiries when I asked them". These findings lead us to the conclusion that GAIER can be more advantageous to students than a conventional course.

Yang *et al. Int J Educ Technol High Educ*   (2023) 20:18

Page 24 of 31

Discussion and conclusions

Discussion

In the past, students' learning interest in traditional laboratory safety courses was low and the uncertainty of teaching made the teaching workload high, while the gamified AIER not only enhances students' learning motivation, but also exists as a teacher's role and reduces the teacher's burden. Therefore, this study developed a gamified AIER learning system with the GAFCC model for laboratory safety courses to investigate its effects on students' learning achievement, cognitive load, learning motivation, flow experience, problem-solving tendency, and student learning behaviors.

In response to research questions 1 and 2, the results showed that students who used a gamified AIER system with the GAFCC model and those who used a general AR system achieved better learning achievement and lower cognitive load than students who learned through traditional instruction. This may be because the use of robots for teaching and learning enhances students' active learning, and timely feedback mechanisms help learners understand their mastery of knowledge and reflect on it. This is in line with the findings of Huang et al. (2019). When students learn with the gamified AIER, they are given tasks in the learning system and informed that they can earn "badges" and "tickets" for success, which facilitate their retention of the learning materials. In their study, Chu et al. (2022b) suggest that learning achievement measurement is highly interested in robotics research projects from a cognitive perspective. In short, robots help students become more motivated and provide sufficient feedback during the learning process to facilitate student learning (Cheniti Belcadhi, 2016; Essel et al., 2022).

Students in the GAIER and AR groups had a lower cognitive load than those in the TI group, probably because the instruction-using robots were already able to encourage active learning among the students and reduce their cognitive loads. Student 3 in the GAIER group's interviews stated, for instance, that "the robot was more interactive than traditional teaching; the real-time feedback from the test made me feel involved in the classroom; and the robot can help me answer questions and give more concise information that I can understand". Similarly, students in the AR group's interviews felt that the robot was very useful because it could give timely support when she did not understand questions.

To answer research questions 3 and 4, the experimental results also showed that using gamified AIER with the GAFCC model motivated students to learn so as to have a better flow experience than the general AR and traditional instruction. The GAIER group employed the GAFCC model, a gamification model built based on the five motivation theories. And the five-stage gamification design technique also provided information on how the GAFCC model was to be applied to the classroom in a more intentional way. The gamification and GAFCC models, on the other hand, were not explored by either the AR group or the traditional group. Additionally, one student from the AR group stated in the interviews that "robotics was quite exciting the first time I was introduced to it, but I lost interest after I learned it, and I would probably not like it if it were taught in this way again." However, the majority of the students in the GAIER group were eager to utilize the system once more and advocate for it to their classmates and instructors. The GAIER group, meanwhile, generated the lesson plans using the gamification strategy of the GAFCC model, which encouraged students' learning from five dimensions,

including cooperative and competitive activities to promote students' active learning, the combination of unlocking levels and task progress to improve students' learning behavior, and the distinction of different level settings to bring students certain novelty and immersion. In this way, the GAIER group outperformed both the AR and traditional groups in terms of learning motivation and flow experiences. These results are consistent with those of Lee et al. (2022), who investigated the effects of AI chatbots on college students' learning motivation and attitudes. Both of these factors helped students become more motivated in their studies and enhance their learning motivation.

To answer research question 5, the results showed that, compared to traditional instruction, the gamified AIER system and the general AR system greatly increased students' tendency for problem-solving. One the one hand, the GAIER group had staged tasks, cooperation, and obstacles, while the AR group has more cooperative answers to questions. This helped both groups of students actively ask the robot for learning materials and then complete the test. Some studies have shown that effective peer interaction can help them identify learning problems and possible ways to find solutions (Merrill & Gilbert, 2008). On the other hand, feedback mechanisms were set up in both the GAIER and the AR groups. After students responded, the system immediately provided feedback and references, and students could also see them repeated times. As a result, both the GAIER and AR groups showed higher problem-solving tendency than the TI group, and this result is consistent with the findings of Hwang et al. (2014), which concluded that both effective peer collaboration and feedback mechanisms can promote students' problem-solving skills.

To answer research question 6, differences in learning behaviors between the GAIER and AI groups were analyzed. The results showed that the students in the GAIER group were more able to bring into the learning situation than the students in the AR group during the learning process. The behavioral pattern of the two groups showed that the timely feedback given by the robot is also a key function of the learning process (Epstein et al., 2002), which promotes active learning and motivates students to participate in learning to achieve their learning goals. It was discovered that L1 → P1 behavior was significantly different between the GAIER and AR groups. There are two possible explanations for this: first, the extensive learning content necessitated that students repeatedly ask the robot questions in order to obtain learning materials and then learn them; and second, the cooperation process was poorly coordinated, resulting in only one party interacting with the robot while the other party watched, which resulted in individual students' unrelated learning behaviors. The study's findings, however, demonstrated that the inclusion of gamified components helped students focus on their work while they were learning and kept them from becoming overly bored for extended periods of time. The learning materials and content were more comprehensive in the GAIER group, and the students displayed more concentrated and engaged emotions. Students in the AR group, on the other hand, initially expressed enthusiasm in interacting with the robot, but this interest gradually waned and there were little emotional ups and downs.

Conclusions and suggestions

In conclusion, gamified AIER significantly improved students' learning achievement, learning motivation, flow experience, and problem-solving tendency, and also reduced

the cognitive load of learning. Additionally, students have better behavior patterns when supported by gamification strategies based on the GAFCC model. The five-stage instructional design model is more relevant to the learning process of the students, integrating them into the learning environment, and making it easier for them to accomplish the learning activities, and enjoy the learning experience. GAIER is a computer-simulated teacher who engages with students through specified information, assists them in carrying out their studies, assesses their understanding, and provides feedback as they learn. Additionally, during the interaction process, GAIER was designed to provide students with vocal, physical, visual, and textual encouragement in the form of emojis, victory sound effects, points, and badges. Students' motivation and flow experiences are nurtured in this way, which enhances their capacity for learning. "GAIER captures my interest and then engages me more in studying, and the interaction allows me to delve deeper into my knowledge," remarked S101 from the GAIER group.

From the perspective of educational system development, the current study further reveals the need to incorporate appropriate strategies in the development of AIER to enable active and immersive student participation in learning. Gamification strategies based on the GAFCC model enhance students' learning experience with the robot, prompting learners to show joy and excitement in the learning process. From a developmental perspective, AI is the trend of development, and incorporating AR with interactive, speech systems into the classroom can promote active student exploration and problem-solving tendency, providing different options for future research in AIER to explore. From a subject perspective, the teaching style is innovative because it uses a robot as the only teacher to lead students through the learning process. Teachers might avoid repeating the same information by merely summarizing and emphasizing the key points at the lesson's conclusion. This allows students to learn in a novel way. In the interviews, S105 of the GAIER group believed that first aid knowledge in medical education is suitable to be coupled with such robots, S205 of the AR group thought it was suitable for science knowledge, and S207 said that boring literature knowledge could be combined with robots. As a result, GAIER can be used to teach some necessary, timely, and repetitive knowledge. From the perspective of the appropriate users, college students are the study's target group since they are more adept at utilizing electronic devices, have a certain amount of background knowledge, and can grab robot operation more quickly. However, it will be challenging for elementary and middle school kids to reach the trial outcomes, and a teacher will be required to help them. Therefore, high school, college, and vocational students are the ideal options for the robot's function as a teacher because they have a particular level of knowledge and are technologically savvy.

Also, there are currently some limitations to this study that need to be noted. First, due to the small sample size, the results cannot be generalized to students' learning in all situations. Second, the coding content of the behavioral sequence analysis was not detailed enough, and in order to ensure that the students' actions could be seen, the recording with a high-definition camera mainly captured the students' side faces and the physical activities on the upper itself, and the students' expressions could not be seen head-on, making it impossible for the researcher to observe whether the students showed any dazed or closed-eye behaviors while reading. Finally, the game is very time-consuming to design, develop, and apply. To address the above limitations, this paper makes several

suggestions. First, increase the experimental time and collect more abundant and comprehensive experimental data to prove the effectiveness of the proposed method. Second, in terms of coding content, dual-camera recording can be considered to analyze students' facial expressions from multiple angles, and student behavior can also be analyzed based on big data. For example, by writing codes to capture students' facial emotions using artificial intelligence technology, big data calculates the frequency of students' facial expression changes and categorizes them, avoiding the shortcomings of unclear recognition by the naked eye and further exploring students' behavior patterns. Third, the system development incorporates the GAFCC model based on gamification, and subsequent development can be continued by educators based on the model if the system supports game mechanics. Much of the success of gamification in this study was attributed to following the combination of the GAFCC model and the anthropomorphic robot, and future research could be tested on the basis of an iterative GAFCC model combined with different classes of robots.

Declarations

References

Alimisis, D. (2009). School of Pedagogical and Technological Education (Athens), & TERECoP Project. *Teacher education on robotics: enhanced constructivist pedagogical methods*. School of Pedagogical and Technological Education.

Ayi, H.-R., & Hon, C.-Y. (2018). Safety culture and safety compliance in academic laboratories: A Canadian perspective. *Journal of Chemical Health and Safety, 25*(6), 6–12. https://doi.org/10.1016/j.jchas.2018.05.002

Bai, M., Liu, Y., Qi, M., Roy, N., Shu, C.-M., Khan, F., & Zhao, D. (2022b). Current status, challenges, and future directions of university laboratory safety in China. *Journal of Loss Prevention in the Process Industries, 74*, 104671. https://doi.org/10.1016/j.jlp.2021.104671

Bai, S., Hew, K. F., Gonda, D. E., Huang, B., & Liang, X. (2022a). Incorporating fantasy into gamification promotes student learning and quality of online interaction. *International Journal of Educational Technology in Higher Education, 19*(1), 29. https://doi.org/10.1186/s41239-022-00335-9

Bakeman, R., & Gottman, J. M. (1997). *Observing interaction: An introduction to sequential analysis*. Cambridge: Cambridge University Press.

Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. In *Sci. Robot* (Vol. 3). https://www.science.org

Benitti, F. B. V. (2012). Exploring the educational potential of robotics in schools: A systematic review. *Computers & Education, 58*(3), 978–988. https://doi.org/10.1016/j.compedu.2011.10.006

Benkerzaz, S., Elmir, Y., & Dennai, A. (2019). A Study on Automatic Speech Recognition saliha 08 A Study on automatic speech recognition a study on automatic speech recognition. *Journal of Information Technology Review, 10*. https://doi.org/10.6025/jitr/2019/10/3/77-85.

Bers, M. U., Flannery, L., Kazakoff, E. R., & Sullivan, A. (2014). Computational thinking and tinkering: Exploration of an early childhood robotics curriculum. *Computers and Education, 72*, 145–157. https://doi.org/10.1016/j.compedu.2013.10.020

Blanchard, S., Freiman, V., & Lirrete-Pitre, N. (2010). Strategies used by elementary schoolchildren solving robotics-based complex tasks: Innovative potential of technology. *Procedia Social and Behavioral Sciences, 2*(2), 2851–2857. https://doi.org/10.1016/j.sbspro.2010.03.427

Chandra, S., Dillenbourg, P., & Paiva, A. (2020). Children teach handwriting to a social robot with different learning competencies. *International Journal of Social Robotics, 12*(3), 721–748. https://doi.org/10.1007/s12369-019-00589-w

Chen, H., Park, H. W., & Breazeal, C. (2020c). Teaching and learning with children: Impact of reciprocal peer learning with a social robot on children's learning and emotive engagement. *Computers and Education, 150*, 103836. https://doi.org/10.1016/j.compedu.2020.103836

Chen, L., Chen, P., & Lin, Z. (2020a). Artificial intelligence in education: A review. *IEEE Access, 8*, 75264–75278. https://doi.org/10.1109/ACCESS.2020.2988510

Chen, X., Xie, H., Zou, D., & Hwang, G.-J. (2020b). Application and theory gaps during the rise of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence, 1*, 100002. https://doi.org/10.1016/j.caeai.2020.100002

ChenitiBelcadhi, L. (2016). Personalized feedback for self assessment in lifelong learning environments based on semantic web. *Computers in Human Behavior, 55*, 562–570. https://doi.org/10.1016/j.chb.2015.07.042

Chevalier, M., Giang, C., Piatti, A., & Mondada, F. (2020). Fostering computational thinking through educational robotics: A model for creative computational problem solving. *International Journal of STEM Education, 7*(1), 39. https://doi.org/10.1186/s40594-020-00238-z

Chin, K.-Y., Wu, C.-H., & Hong, Z.-W. (2011). A humanoid robot as a teaching assistant for primary education. *Fifth International Conference on Genetic and Evolutionary Computing, 2011*, 21–24. https://doi.org/10.1109/ICGEC.2011.13

Chu, S., Hwang, G., Chien, S., & Chang, S. (2022a). Incorporating teacher intelligence into digital games: An expert system-guided self-regulated learning approach to promoting EFL students' performance in digital gaming contexts. *British Journal of Educational Technology*. https://doi.org/10.1111/bjet.13260

Chu, S.-T., Hwang, G.-J., & Tu, Y.-F. (2022b). Artificial intelligence-based robots in education: A systematic review of selected SSCI publications. *Computers and Education: Artificial Intelligence, 3*, 100091. https://doi.org/10.1016/j.caeai.2022.100091

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*(1), 98–104. https://doi.org/10.1037/0021-9010.78.1.98

Csikszentmihalyi, M. (1978). Intrinsic rewards and emergent motivation. In M. R. Lepper & D. Greene (Eds.), *The hidden costs of reward: New perspectives on the psychology of human motivation* (pp. 205–216). Earlbaum.

de-Marcos, L., Garcia-Lopez, E., & Garcia-Cabot, A. (2016). On the effectiveness of game-like and social approaches in learning: Comparing educational gaming, gamification & social networking. *Computers and Education, 95*, 99–113. https://doi.org/10.1016/j.compedu.2015.12.008

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness. *Proceedings of the 15th International Academic MindTrek Conference on Envisioning Future Media Environments-MindTrek '11*, 9. https://doi.org/10.1145/2181037.2181040

Diago, P. D., González-Calero, J. A., & Yáñez, D. F. (2022). Exploring the development of mental rotation and computational skills in elementary students through educational robotics. *International Journal of Child-Computer Interaction, 32*, 100388. https://doi.org/10.1016/j.ijcci.2021.100388

Domínguez, A., Saenz-de-Navarrete, J., de-Marcos, L., Fernández-Sanz, L., Pagés, C., & Martínez-Herráiz, J.-J. (2013). Gamifying learning experiences: Practical implications and outcomes. *Computers and Education, 63*, 380–392. https://doi.org/10.1016/j.compedu.2012.12.020

Epstein, M. L., Lazarus, A. D., Calvano, T. B., Matthews, K. A., Hendel, R. A., Epstein, B. B., & Brosvic, G. M. (2002). Immediate feedback assessment technique promotes learning and corrects inaccurate first responses. *The Psychological Record, 52*(2), 187–201. https://doi.org/10.1007/BF03395423

Essel, H. B., Vlachopoulos, D., Tachie-Menson, A., Johnson, E. E., & Baah, P. K. (2022). The impact of a virtual teaching assistant (chatbot) on students' learning in Ghanaian higher education. *International Journal of Educational Technology in Higher Education, 19*(1), 57. https://doi.org/10.1186/s41239-022-00362-6

Evripidou, S., Georgiou, K., Doitsidis, L., Amanatiadis, A. A., Zinonos, Z., & Chatzichristofis, S. A. (2020). Educational robotics: platforms, competitions and expected learning outcomes. *IEEE Access, 8*, 219534–219562. https://doi.org/10.1109/ACCESS.2020.3042555

Falkner, N. J. G., & Falkner, K. E. (2014). "Whither, badges?" or "wither, badges!" *Proceedings of the 14th Koli Calling International Conference on Computing Education Research-Koli Calling '14*, 127–135. https://doi.org/10.1145/2674683.2674698

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*(2), 117–140. https://doi.org/10.1177/001872675400700202

Fridin, M. (2014). Storytelling by a kindergarten social assistive robot: A tool for constructive learning in preschool education. *Computers and Education, 70*, 53–64. https://doi.org/10.1016/j.compedu.2013.07.043

Fryer, L. K., Ainley, M., Thompson, A., Gibson, A., & Sherlock, Z. (2017). Stimulating and sustaining interest in a language course: An experimental comparison of Chatbot and Human task partners. *Computers in Human Behavior, 75*, 461–468. https://doi.org/10.1016/j.chb.2017.05.045

Ge, Z.-G. (2018). The impact of a forfeit-or-prize gamified teaching on e-learners' learning performance. *Computers and Education, 126*, 143–152. https://doi.org/10.1016/j.compedu.2018.07.009

Groening, C., & Binnewies, C. (2019). "Achievement unlocked!"—The impact of digital achievements as a gamification element on motivation and performance. *Computers in Human Behavior, 97*, 151–166. https://doi.org/10.1016/j.chb.2019.02.026

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does Gamification Work?—A literature review of empirical studies on gamification. *2014 47th Hawaii International Conference on System Sciences*, 3025–3034. https://doi.org/10.1109/HICSS.2014.377.

Hanus, M. D., & Fox, J. (2015). Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Computers and Education, 80*, 152–161. https://doi.org/10.1016/j.compedu.2014.08.019

Hassan, M. A., Habiba, U., Majeed, F., & Shoaib, M. (2021). 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

Hew, K. F., Huang, B., Chu, K. W. S., & Chiu, D. K. W. (2016). Engaging Asian students through game mechanics: Findings from two experiment studies. *Computers and Education, 92–93*, 221–236. https://doi.org/10.1016/j.compedu.2015.10.010

Hsu, C.-C., & Wang, T.-I. (2018). Applying game mechanics and student-generated questions to an online puzzle-based game learning system to promote algorithmic thinking skills. *Computers and Education, 121*, 73–88. https://doi.org/10.1016/j.compedu.2018.02.002

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 and Education, 125*, 254–272. https://doi.org/10.1016/j.compedu.2018.06.018

Huang, B., & Hew, K. (2021). Using gamification to design courses: lessons learned in a three-year design-based study. *Educational Technology and Society, 24*, 44–63.

Huang, B., Hwang, G.-J., Hew, K. F., & Warning, P. (2019). Effects of gamification on students' online interactive patterns and peer-feedback. *Distance Education, 40*(3), 350–379. https://doi.org/10.1080/01587919.2019.1632168

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

Hwang, G. J., & Chang, C. Y. (2021). *A review of opportunities and challenges of chatbots in education*. Routledge. https://doi.org/10.1080/10494820.2021.1952615

Hwang, G. J., Chang, S. C., Chen, P. Y., & Chen, X. Y. (2018). Effects of integrating an active learning-promoting mechanism into location-based real-world learning environments on students' learning performances and behaviors. *Educational Technology Research and Development, 66*(2), 451–474. https://doi.org/10.1007/s11423-017-9567-5

Hwang, G.-J., Hung, C.-M., & Chen, N.-S. (2014). Improving learning achievements, motivations and problem-solving skills through a peer assessment-based game development approach. *Educational Technology Research and Development, 62*(2), 129–145. https://doi.org/10.1007/s11423-013-9320-7

Hwang, G.-J., Sung, H.-Y., Chang, S.-C., & Huang, X.-C. (2020). A fuzzy expert system-based adaptive learning approach to improving students' learning performances by considering affective and cognitive factors. *Computers and Education: Artificial Intelligence, 1*, 100003. https://doi.org/10.1016/j.caeai.2020.100003

Hwang, G.-J., Yang, L.-H., & Wang, S.-Y. (2013). A concept map-embedded educational computer game for improving students' learning performance in natural science courses. *Computers and Education, 69*, 121–130. https://doi.org/10.1016/j.compedu.2013.07.008

Hwang, G. J., Yang, T. C., Tsai, C. C., & Yang, S. J. H. (2009). A context-aware ubiquitous learning environment for conducting complex science experiments. *Computers and Education, 53*(2), 402–413.

Kalogiannakis, M., Papadakis, S., & Zourmpakis, A. I. (2021). Gamification in science education. A systematic review of the literature. *Education Sciences, 11*(1), 1–36. https://doi.org/10.3390/educsci11010022

Kapp, K. M. (2012). The gamification of learning and instruction: Game-based methods and strategies for training and education. *International Journal of Gaming and Computer-Mediated Simulations, 4*(4), 81–83. https://doi.org/10.4018/jgcms.2012100106

Kasahara, R., Sakamoto, K., Washizaki, H., & Fukazawa, Y. (2019). Applying gamification to motivate students to write high-quality code in programming assignments. *Proceedings of the 2019 ACM Conference on Innovation and Technology in Computer Science Education*, 92–98. https://doi.org/10.1145/3304221.3319792

Kasap, Z., & Magnenat-Thalmann, N. (2012). Building long-term relationships with virtual and robotic characters: The role of remembering. *The Visual Computer, 28*(1), 87–97. https://doi.org/10.1007/s00371-011-0630-7

Kelleher, C., Pausch, R., & Kiesler, S. (2007). Storytelling alice motivates middle school girls to learn computer programming. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1455–1464. https://doi.org/10.1145/1240624.1240844

Kewalramani, S., Kidman, G., & Palaiologou, I. (2021). Using artificial intelligence (AI)-interfaced robotic toys in early childhood settings: A case for children's inquiry literacy. *European Early Childhood Education Research Journal, 29*(5), 652–668. https://doi.org/10.1080/1350293X.2021.1968458

Kyewski, E., & Krämer, N. C. (2018). To gamify or not to gamify? An experimental field study of the influence of badges on motivation, activity, and performance in an online learning course. *Computers and Education, 118*, 25–37. https://doi.org/10.1016/j.compedu.2017.11.006

Lai, C. L., & Hwang, G. J. (2014). Effects of mobile learning participation time on high school students' 21st century core competences. *Proceedings—2014 International Conference of Educational Innovation Through Technology, EITT 2014*, 205–211. https://doi.org/10.1109/EITT.2014.40

Lee, Y.-F., Hwang, G.-J., & Chen, P-Y. (2022). Impacts of an AI-based chabot on college students' after-class review, academic performance, self-efficacy, learning attitude, and motivation. *Educational Technology Research and Development*. https://doi.org/10.1007/s11423-022-10142-8

Lewis Johnson, W. (2001). *Pedagogical Agent Research at CARTE*.

Leyzberg, D., Spaulding, S., Toneva, M., & Scassellati, B. (2012). The physical presence of a robot tutor increases cognitive learning gains. *Proceedings of the Annual Meeting of the Cognitive Science Society, 34*, 34.

Limna, P., Jakwatanatham, S., Siripipattanakul, S., Kaewpuang, P., & Sriboonruang, P. (2022). *A Review of Artificial Intelligence (AI) in Education during the Digital Era*. https://ssrn.com/abstract=4160798

Lin, V., Yeh, H.-C., & Chen, N.-S. (2022). A systematic review on oral interactions in robot-assisted language learning. *Electronics, 11*(2), 290. https://doi.org/10.3390/electronics11020290

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist, 57*(9), 705–717. https://doi.org/10.1037/0003-066X.57.9.705

Maximova, M., & Kim, Y. (2016). *Research trend analysis on the usage of robotics in education. 12*(1), 45–60. https://doi.org/10.14580/apcj.2016.12.1.45

Ménard, A. D., & Trant, J. F. (2020). A review and critique of academic lab safety research. *Nature Chemistry, 12*(1), 17–25. https://doi.org/10.1038/s41557-019-0375-x

Merrill, M. D., & Gilbert, C. G. (2008). Effective peer interaction in a problem-centered instructional strategy. *Distance Education, 29*(2), 199–207. https://doi.org/10.1080/01587910802154996

Muniasamy, A., & Alasiry, A. (2020). Deep learning: The impact on future eLearning. *International Journal of Emerging Technologies in Learning (IJET), 15*(01), 188. https://doi.org/10.3991/ijet.v15i01.11435

Nadi-Ravandi, S., & Batooli, Z. (2022). Gamification in education: A scientometric, content and co-occurrence analysis of systematic review and meta-analysis articles. *Education and Information Technologies, 27*(7), 10207–10238. https://doi.org/10.1007/s10639-022-11048-x

Osatuyi, B., Osatuyi, T., & de la Rosa, R. (2018). Systematic review of gamification research in IS education: A multi-method approach. *Communications of the Association for Information Systems*, *42*. https://doi.org/10.17705/1CAIS.04205

Papert, S. (1980). *Mindstorms: children, computers, and powerful ideas*.

Pearce, J. M., Ainley, M., & Howard, S. (2005). The ebb and flow of online learning. *Computers in Human Behavior, 21*(5), 745–771. https://doi.org/10.1016/j.chb.2004.02.019

Pei, Z., & Nie, Y. (2018). Educational robots: Classification, characteristics, application areas and problems. *Seventh International Conference of Educational Innovation through Technology (EITT), 2018*, 57–62. https://doi.org/10.1109/EITT.2018.00020

Pintrich, P.R., Smith, D.A.F., Garcia, T. and McKeachie, W.J. (1991) *A Manual for the Use of the Motivated Strategies for Learning Questionnaire,* Tech. Rep. No. 91-B-004, University of Michigan, Ann Arbor, MI.

Quera, V., Bakeman, R., & Gnisci, A. (2007). Observer agreement for event sequences: Methods and software for sequence alignment and reliability estimates. *Behavior Research Methods, 39*(1), 39–49. https://doi.org/10.3758/BF03192842

Reyes, G. E. B., López, E., Ponce, P., & Mazón, N. (2021). Role assignment analysis of an assistive robotic platform in a high school mathematics class, through a gamification and usability evaluation. *International Journal of Social Robotics, 13*(5), 1063–1078. https://doi.org/10.1007/s12369-020-00698-x

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*(1), 54–67. https://doi.org/10.1006/ceps.1999.1020

Sailer, M., Hense, J. U., Mayr, S. K., & Mandl, H. (2017). How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in Human Behavior, 69*, 371–380. https://doi.org/10.1016/j.chb.2016.12.033

Salkind, N. J. (2017). *Tests & measurement for people who (think they) hate tests & measurement*. Sage.

Scaradozzi, D., Screpanti, L., & Cesaretti, L. (2019). Towards a definition of educational robotics: A classification of tools experiences and assessments. *Smart learning with educational robotics* (pp. 63–92). Springer Internation Publishing. https://doi.org/10.1007/978-3-030-19913-5_3

Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action: A survey. *International Journal of Human–computer Studies, 74*, 14–31. https://doi.org/10.1016/j.ijhcs.2014.09.006

Shorey, S., Ang, E., Yap, J., Ng, E. D., Lau, S. T., & Chui, C. K. (2019). A virtual counseling application using artificial intelligence for communication skills training in nursing education: development study. *Journal of Medical Internet Research, 21*(10), e14658. https://doi.org/10.2196/14658

Skinner, B. F. (1953). *Science and human behavior*. Macmillan.

Smutny, P., & Schreiberova, P. (2020). Chatbots for learning: A review of educational chatbots for the Facebook Messenger. *Computers and Education, 151*, 103862. https://doi.org/10.1016/j.compedu.2020.103862

Sophokleous, A., Christodoulou, P., Doitsidis, L., & Chatzichristofis, S. A. (2021). Computer vision meets educational robotics. *Electronics, 10*(6), 730. https://doi.org/10.3390/electronics10060730

Subhash, S., & Cudney, E. A. (2018). Gamified learning in higher education: A systematic review of the literature. *Computers in Human Behavior, 87*, 192–206. https://doi.org/10.1016/j.chb.2018.05.028

Sullivan, A., & Bers, M. U. (2018). Dancing robots: Integrating art, music, and robotics in Singapore's early childhood centers. *International Journal of Technology and Design Education, 28*(2), 325–346. https://doi.org/10.1007/s10798-017-9397-0

Sun, J., Gu, C., Chen, J., Wei, W., Yang, C., & Jiang, Q. (2022). *A study of the effects of interactive AI image processing functions on children's painting education* (pp. 93–108). Springer. https://doi.org/10.1007/978-3-031-05581-2_8

Sung, H.-Y., & Hwang, G.-J. (2013). A collaborative game-based learning approach to improving students' learning performance in science courses. *Computers and Education, 63*, 43–51. https://doi.org/10.1016/j.compedu.2012.11.019

Sung, H. Y., & Hwang, G. J. (2018). Facilitating effective digital game-based learning behaviors and learning performances of students based on a collaborative knowledge construction strategy. *Interactive Learning Environments, 26*(1), 118–134. https://doi.org/10.1080/10494820.2017.1283334

Tanaka, F., & Matsuzoe, S. (2012). Children teach a care-receiving robot to promote their learning: field experiments in a classroom for vocabulary learning. *Journal of Human-Robot Interaction*, 78–95. https://doi.org/10.5898/JHRI.1.1.Tanaka

Timms, M. J. (2016). Letting artificial intelligence in education out of the box: educational cobots and smart classrooms. *International Journal of Artificial Intelligence in Education, 26*(2), 701–712. https://doi.org/10.1007/s40593-016-0095-y

Viitaharju, P., Yliniemi, K., Nieminen, M., & Karttunen, A. J. (2021). Learning experiences from digital laboratory safety training. *Education for Chemical Engineers, 34*, 87–93. https://doi.org/10.1016/j.ece.2020.11.009

Wallace, S. A., McCartney, R., & Russell, I. (2010). Games and machine learning: A powerful combination in an artificial intelligence course. *Computer Science Education, 20*(1), 17–36. https://doi.org/10.1080/08993400903525099

Wang, L., & Chen, M. (2010). The effects of game strategy and preference-matching on flow experience and programming performance in game-based learning. *Innovations in Education and Teaching International, 47*(1), 39–52. https://doi.org/10.1080/14703290903525838

Wu, E.H.-K., Lin, C.-H., Ou, Y.-Y., Liu, C.-Z., Wang, W.-K., & Chao, C.-Y. (2020). Advantages and constraints of a hybrid model K-12 E-learning assistant chatbot. *IEEE Access, 8*, 77788–77801. https://doi.org/10.1109/ACCESS.2020.2988252

Yang, J., & Zhang, B. (2019). Artificial intelligence in intelligent tutoring robots: a systematic review and design guidelines. *Applied Sciences, 9*(10), 2078. https://doi.org/10.3390/app9102078

Yannakakis, G. N., & Togelius, J. (2018). Artificial intelligence and games. *Springer International Publishing*. https://doi.org/10.1007/978-3-319-63519-4

Zainuddin, Z., Chu, S. K. W., Shujahat, M., & Perera, C. J. (2020). The impact of gamification on learning and instruction: A systematic review of empirical evidence. *Educational Research Review, 30*, 100326. https://doi.org/10.1016/j.edurev.2020.100326

Zhang, J. H., Zou, L., Miao, J., Zhang, Y. X., Hwang, G. J., & Zhu, Y. (2020). An individualized intervention approach to improving university students' learning performance and interactive behaviors in a blended learning environment. *Interactive Learning Environments, 28*(2), 231–245. https://doi.org/10.1080/10494820.2019.1636078

Zhang, K., & Aslan, A. B. (2021). AI technologies for education: Recent research & future directions. *Computers and Education: Artificial Intelligence, 2*, 100025. https://doi.org/10.1016/j.caeai.2021.100025

Zhao, J., Hwang, G. J., Chang, S. C., Yang, Q. F., & Nokkaew, A. (2021). Effects of gamified interactive e-books on students' flipped learning performance, motivation, and meta-cognition tendency in a mathematics course. *Educational Technology Research and Development, 69*(6), 3255–3280. https://doi.org/10.1007/s11423-021-10053-0

Zhong, B., & Xia, L. (2020). A systematic review on exploring the potential of educational robotics in mathematics education. *International Journal of Science and Mathematics Education, 18*(1), 79–101. https://doi.org/10.1007/s10763-018-09939-y

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.