Received: 23 March 2024 | Accepted: 28 September 2024

DOI: 10.1111/bjet.13526

ORIGINAL ARTICLE

British Journal of Educational Technology BERA

Enhancing student experience and learning with iterative design in an intelligent educational game

Ryan Hare[1] | Sarah Ferguson[2] | Ying Tang[1]

[1]Department of Electrical and Computer Engineering, Rowan University, Glassboro, New Jersey, USA

[2]Department of Education, Rowan University, Glassboro, New Jersey, USA

Correspondence
Ryan Hare, Department of Electrical and Computer Engineering, Rowan University, Glassboro, NJ, USA.
Email: harer6@rowan.edu

Funding information
National Science Foundation, Grant/Award Number: 1913809

With increasing interest in computer-assisted education, AI-integrated systems become highly applicable with their ability to adapt based on user interactions. In this context, this paper focuses on understanding and analysing first-year undergraduate student responses to an intelligent educational system that applies multi-agent reinforcement learning as an AI tutor. With human–computer interaction at the centre, we discuss principles of interface design and educational gamification in the context of multiple years of student observations, student feedback surveys and focus group interviews. We show positive feedback from the design methodology we discuss as well as the overall process of providing automated tutoring in a gamified virtual environment. We also discuss students' thinking in the context of gamified educational systems, as well as unexpected issues that may arise when implementing such systems. Ultimately, our design iterations and analysis both offer new insights for practical implementation of computer-assisted educational systems, focusing on how AI can augment, rather than replace, human intelligence in the classroom.

KEYWORDS
computer-assisted education, human–computer interaction, serious games

Practitioner notes

What is already known about this topic

- AI-integrated systems show promise for personalizing learning and improving student education.
- Existing research has shown the value of personalized learner feedback.
- Engaged students learn more effectively.

What this paper adds

- Student opinions of and responses to an HCI-based personalized educational system.
- New insights for practical implementation of AI-integrated educational systems informed by years of student observations and system improvements.
- Qualitative insights into system design to improve human–computer interaction in educational systems.

Implications for practice and/or policy

- Actionable design principles for computer-assisted tutoring systems derived from first-hand student feedback and observations.
- Encourage new directions for human–computer interaction in educational systems.

INTRODUCTION

With modern AI approaches, education faces a transformative era in terms of the facilitation and personalization of learning (Kumar et al., 2023). By helping to automate traditional teaching or providing analytics on learner performance (Alloghani et al., 2018; Anantharaman et al., 2018), new AI methodologies continue to push education forward. Using AI, educational systems can be made more automated, providing more effective (Chaudhry & Kazim, 2022) and personalized educational experiences to students (Limna et al., 2022), often through intelligent tutoring systems (ITSs) (Lin et al., 2023). ITSs leverage the power of human–computer interaction (HCI) to create or enhance students' classroom experiences beyond what is typical. However, the successful implementation of such systems requires both advanced AI methodologies and careful attention to the HCI aspects, ensuring that integration enhances the learning process without any negative effects on students' learning.

When implementing AI-assisted educational systems, research in educational technology consistently highlights the importance of both personalized feedback (Zheng et al., 2022) and learner engagement (Bond & Bedenlier, 2019) when trying to effectively deliver educational content and experiences. However, many unforeseen challenges can arise when implementing such systems in a real classroom setting, and controlled lab environments can often fail to reveal potential issues. Furthermore, learner engagement and effective education are often heavily dependent on user perceptions, which can fail to align with designer intent. This work explores both possible unexpected challenges in a classroom implementation and how learner feedback can be used to inform better HCI design in an educational system. With that exploration in mind, this paper aims to answer the following research questions:

RQ1. What central aspects of educational system design are important to improve the system's usability and utility to students and ultimately enhance students' learning experience?
RQ2. What unexpected or uncontrollable issues interfere with students' learning during implementation of an educational system?
RQ3. Does an iterative design process informed by student questionnaires lead to improved student perceptions about system usability and utility?

In pursuit of these questions, we provide design insights for effective HCI in classroom settings informed by the implementation and testing of an AI-driven gamified ITS. We focus on student survey responses and results from personal focus group interviews where students detailed their thought processes and opinions on system design, all gathered during a design-based research study. Participating students were all first-year university students in an engineering program, with the serious game in question created as a gamified lab experience in a foundational course. Through our qualitative analysis, we provide other researchers with important design principles and considerations when implementing this type of system, all of which may not be obvious in a study focused purely on quantitative analysis. We focus on the nuance of human–AI interaction and detail how such systems can be designed and implemented to maximize not just the educational benefits but also the learner experience. And ultimately, we aim to provide insights from our development to inform other researchers' designs and implementations.

RELATED LITERATURE

Research in HCI is heavily intertwined with the development of educational technology, as the way users interact with a system directly influences its effectiveness as an educational tool. Therefore, HCI research emphasizes the importance of user-centred design (Hasani et al., 2020) to ensure widespread usability for educational systems (Cruz et al., 2015; González-González et al., 2015), as the target audience in educational settings often varies greatly in their prior knowledge, past experiences and expectations about their interactions with such systems (Yi et al., 2021). Ultimately, if educational technology aims to supplement human-driven instruction in any meaningful way, the design aspects surrounding the user's experience and interfacing must not interfere with the delivery of educational content, as poor user interface design can greatly impact system usability (Miya & Govender, 2022) and user retention (Priyadarshini, 2024). Furthermore, education-specific systems often suffer from poor integration with their target curricula (Kalmpourtzis & Romero, 2020).

A key challenge in this area is the discrepancy between designers' visions surrounding a system's use and how users interact with it (Duvaud et al., 2021). Designers, especially those with heavy engineering backgrounds, tend to prioritize technical efficiency and functionality over usability or minute details of the user experience. As a result, educational software often encounters challenges with interface design, user friendliness and usability (Ahmad Faudzi et al., 2023; Faghih et al., 2013). In fact, oversight in usability can cause significant issues with the software's effectiveness as an in-classroom educational tool. Furthermore, when design teams fail to properly test and receive feedback from students, system use can fail to align with designer intent, potentially leading to poor or incorrect delivery of educational content, which could harm students in the long term. With those challenges in mind, this paper details an iterative, feedback-driven approach to our design of a gamified educational tool, focusing on how users' experiences and perceptions can be used to inform design changes that improve the overall educational experience. Our insights into and details around our iterative design process can inform future researchers regarding various design nuances that improve the effectiveness and usability of educational software.

ITERATIVE DESIGN

This paper details our design process and student feedback from the implementation and testing of an AI-driven gamified educational system. Ultimately, our insights presented here aim to inform researchers on future projects about proper design processes, as well as how students react to specific aspects of gamified educational systems. With those goals in mind, this section details the principles and process for our iterative design, the various changes made to our gamified education system, how student feedback informed those changes and students' resulting feedback after said changes were implemented.

Logic model for educational serious games

First and foremost, to help detail our design process, we have encapsulated our focus and ideas within a logic model shown in Figure 1. This logic model provides a structured approach to our planning, implementation and evaluation all throughout our iterative design process.

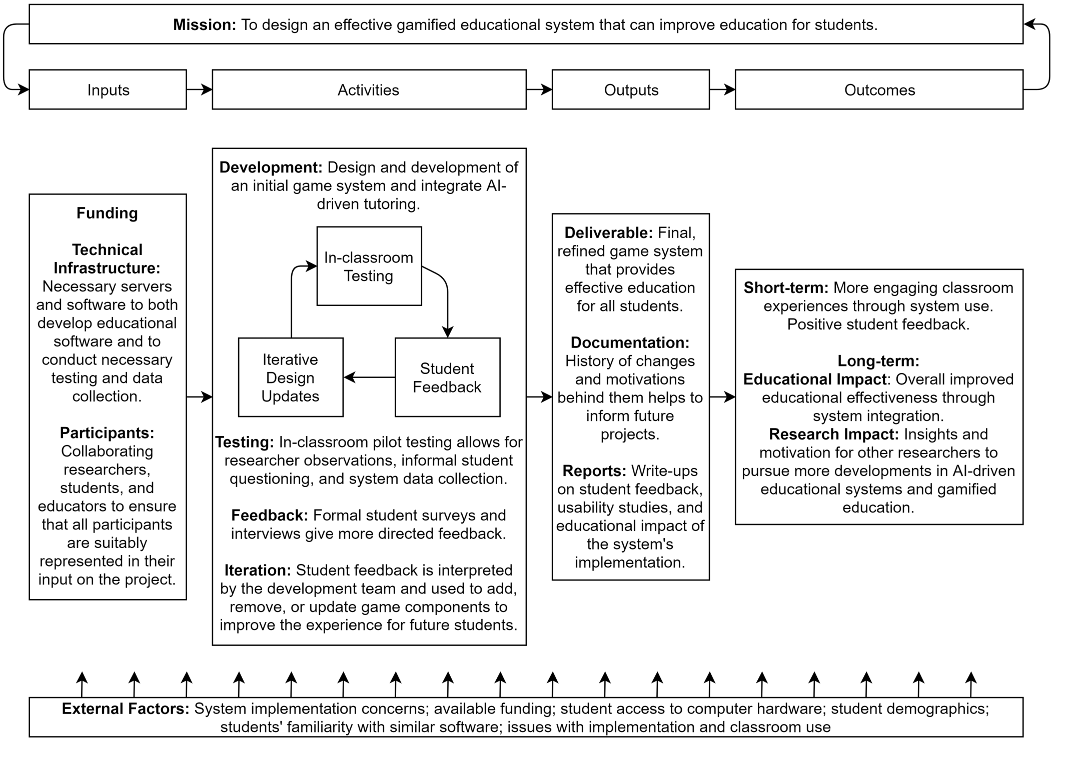

FIGURE 1 Logic model detailing our iterative design process, relevant components, stakeholders and external factors.

Feedback-driven design improvements

With the aforementioned design process in mind, we now discuss issues that we faced in early development and the changes that were made as a result of those issues. With feedback from earlier student testing, these developments aim to answer RQ1 and highlight key aspects of the system's design that helped to improve the system's usability and utility to students.

Game flow

First, to give a more general overview, our system covers content and materials for a first-year introductory electrical engineering course. An accompanying lab assignment tasks students with coding a traffic light, and the game system shares the same end goal, but with varied content delivery. The game intrinsically motivates students at the start by showing a simulated traffic accident caused by a faulty traffic light. In this way, students are encouraged to succeed on their lab project through a "real-world" context, solidifying the problem as a "real" issue instead of just a lab assignment (King & Ritchie, 2012). This intrinsic motivation was present from the initial game version onwards.

As students play, they learn and reinforce necessary background concepts aided by their interactions with the AI tutor, discussed in section "AI-driven student support." The topics presented are subdivided to allow for per-topic coverage and evaluation, giving students necessary background information before they attempt their traffic light design. By scaffolding information, the system ensures that students build solid foundational knowledge before proceeding to solve the overall problem (Puntambekar, 2022). And ultimately, students then submit their traffic light design, which the system automatically checks for functionality.

As in most AI-driven tutoring, our system requires measurements of student performance. At intermittent points in the game, we evaluate students' knowledge on the presented content. Intermittent evaluations help to provide feedback on students' own knowledge, fostering metacognition (Stanton et al., 2021) and encouraging self-reflection and self-assessment (Andrade, 2019). As the game has no failure state, students can also retry evaluation measures as needed. In this repetition, students can reinforce their own concept knowledge (Franzwa et al., 2014). Furthermore, between attempts the AI system provides aid to students, further reinforcing the presented concepts.

When discussing usability and utility for students, evaluation emerged quite early on as a key issue. For instance, Figure 2 shows an early version of the built-in evaluation metrics—multiple choice quizzes. Even early on, students often commented that quizzes and

(a) (b)

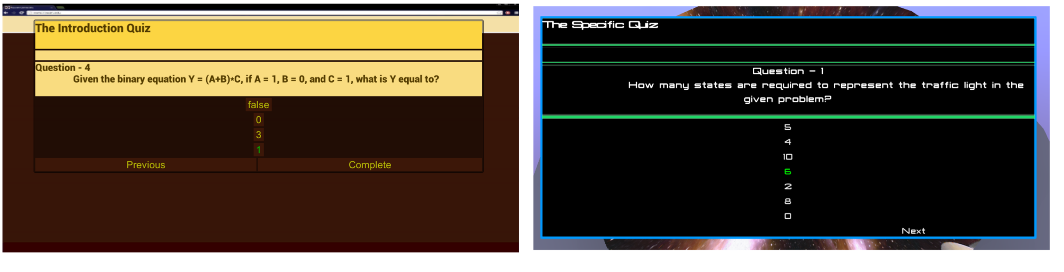

FIGURE 2 Two example images from quizzes in our initial design. These quizzes show (a) a basic question on binary logic and (b) a question regarding the expected traffic light design that students are meant to complete.

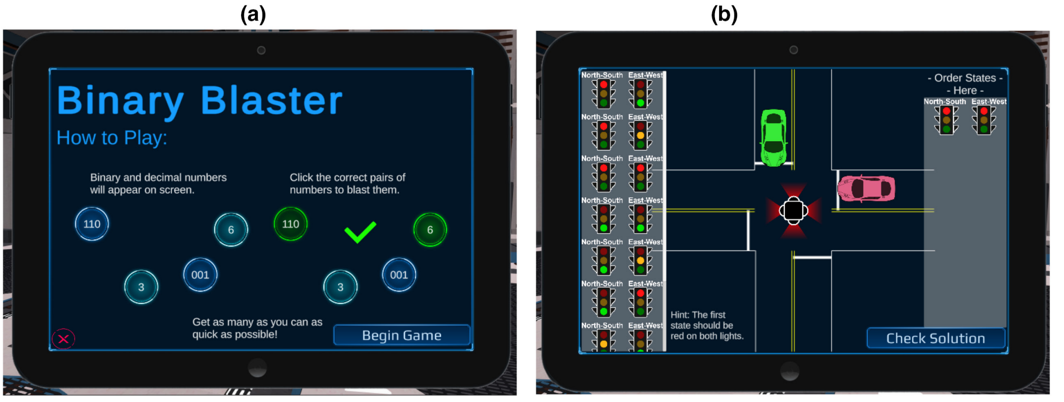

FIGURE 3 Two of the subgames contained within our refined system. (a). "Binary Blaster," a subgame for testing students' knowledge of binary numbers. (b) Two-dimensional traffic light simulator, a subgame for testing students' knowledge of traffic light state ordering.

more overt methods of evaluation detract from the overall experience. Students expect a game to be enjoyable, and quizzes break their flow of play. Furthermore, quizzes are not appropriate for some topics, with Figure 2b showing a question that simply asks students to memorize their lab assignment. Finally, as the game is designed not to have a failure state, students had as many attempts as needed, often leading them to simply memorize answers. Considering Bloom's taxonomy of education, memorization is the lowest order skill in lesson design (Anderson et al., 2000), so encouraging students to apply their knowledge can lead to a deeper and more complete lesson.

Based on the feedback received, new game versions instead gamified the assessment; that is, assessing students' knowledge through puzzles and subgames. Figure 3 shows a few examples. Figure 3a shows "Binary Blaster," a game that tests students' ability to convert between binary and decimal numbers. In this game, students must quickly connect matching numbers, applying their knowledge of how this conversion is normally calculated. Figure 3b shows a two-dimensional traffic light simulator that tests students on their ability to properly order the states of the traffic light. This subgame is especially notable when compared to the quiz question in Figure 2b, which covered the same content, but relied on memorization instead of application of knowledge. In other words, students are encouraged to both know and apply the problem requirements to create a solution to a simplified version of the same problem. In the end, the refined game allows for non-explicit testing with better feedback from students, with students indicating that gamified assessments improve their engagement with the system while allowing us to collect necessary data.

User interface and instruction

A key issue with presenting educational topics through a virtual serious game is students' familiarity with virtual games. Barr (2020) showed that 37% of university students did not play video games (Barr, 2020). So design that might seem intuitive can often be confusing to students who have never touched a virtual game or explored a virtual environment, which many students indicated in our initial version. And ultimately, this impacts students' ability to use the system to its full extent and receive the presented educational material. Based on feedback from students, we focused on addressing this problem, ensuring that goals and necessary interactable items are clearly marked. Furthermore, we added clear in-game documentation on both how students progress and the control scheme used to control the virtual character. Figure 4 shows some examples of instructions presented to students.

For indicating necessary steps to progress through the virtual environment, the design team chose to utilize on-screen waypoint markers shown in Figure 5b, with Figure 5a showing the same game section before waypoints and minor visual changes. In both cases, it is necessary for students to approach the large computer console in the centre, but Figure 5b makes the needed action much clearer to students. Additionally, initial game instructions indicate to students the purpose and appearance of these markers to avoid possible confusion.

In our updated user interface, Figure 6 shows examples of markers indicating recommendations to students and the availability of new content. These markers allow us to give students the freedom to access content in any order if they choose while simultaneously allowing the AI tutoring to recommend certain paths. In this case, students are not forced to engage with help materials but rather are given optional recommendations on how to progress. With this approach, the updated game system can work for self-driven students, allowing them to engage in more self-regulated learning (Sáiz Manzanares et al., 2020) while simultaneously creating a clear and easy path for less motivated or lower-performing students. Ultimately, to ensure the system is usable for students regardless of background, guidance markers and helpful interface elements are key to ensuring that no students get stuck, lost or confused as they progress through the game.

AI-driven student support

For the AI-driven personalization, the system first prompts students with a notification marker like those shown in Figure 6. This allows students to decide whether they want to seek assistance as part of our system's self-regulated nature. When a student does choose to seek assistance, the AI in the system selects appropriate personalized assistance for the student as determined by reinforcement learning.

The full reinforcement learning implementation has been adapted from prior literature (Hare & Tang, 2023) and will be briefly outlined here. As students play, the system records data on their interactions: quiz scores, progression time and actions such as key presses and mouse movements. Students' past data are then used to automatically build a student profile to estimate that student's level of knowledge on a per-topic basis. When

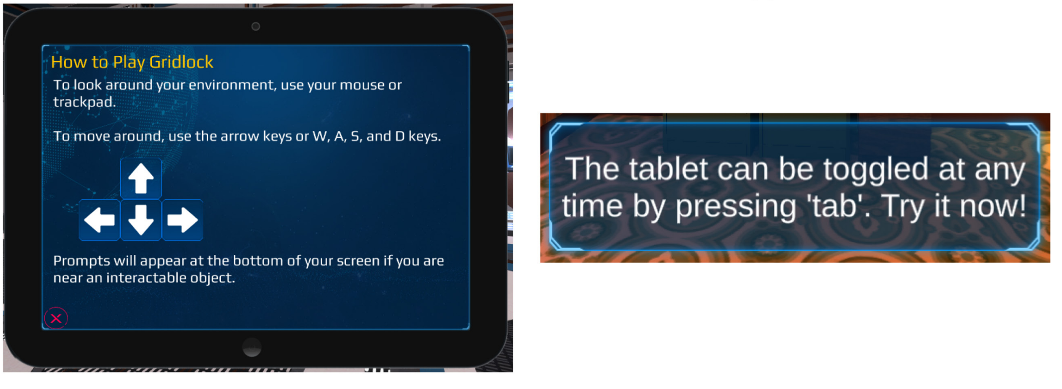

FIGURE 4 Instructions present within the refined game showing students both a) How to play; and b) providing a context-sensitive prompt for interacting with a new menu.

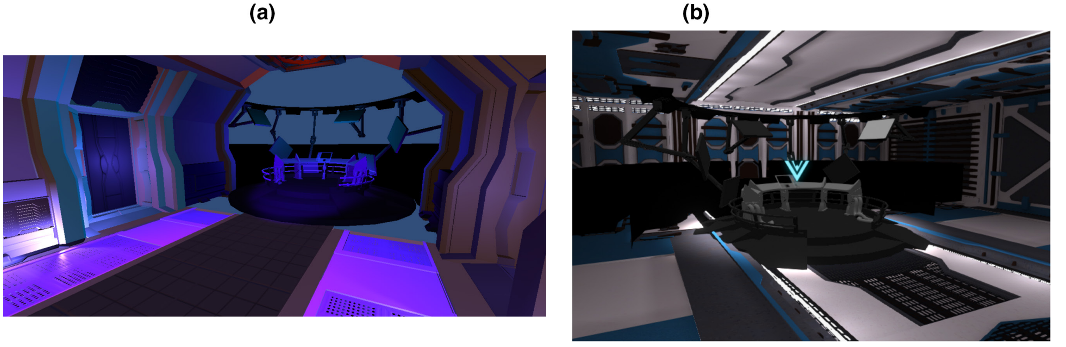

FIGURE 5 (a) Initial game and (b) refined game after adding waypoints. (b) Also includes environmental design changes to help improve the appearance of the virtual environment.

FIGURE 6 Notification markers in the refined game, indicating to students: (a) the "tablet" interface, indicating that there is a new content in the learning module; (b) that assistance on a specific module is recommended; and (c) that viewing a specific module is recommended.

the student seeks assistance, the system can then refer to that student's profile, providing a single numerical response that the game then translates into a pre-made set of study materials. These materials include videos, images, example problems and text, all with varied levels of difficulty and different emotional tones. Then, when the student views this assistance, the game instance observes that student's follow-up performance and uses improvements (or declines) in performance as a numerical score to rate how good the chosen assistance was given that student's initial performance data. With reinforcement learning, this numerical score then informs the system's decisions for future students with similar performance data with the end goal of choosing effective assistance for any and all students. For a more technical overview of this process, we refer readers to the technical literature on the topic (Hare & Tang, 2023).

Figure 7 shows the in-game help documentation before the addition of the AI assistance. The prior document was all-encompassing, attempting to cover every topic for every student, and many students found this overwhelming. Often, students would have to understand what they did not know in order to seek assistance. With the new system, Figure 8 shows a few possible assistance documents. Students might be shown videos (Figure 8a),

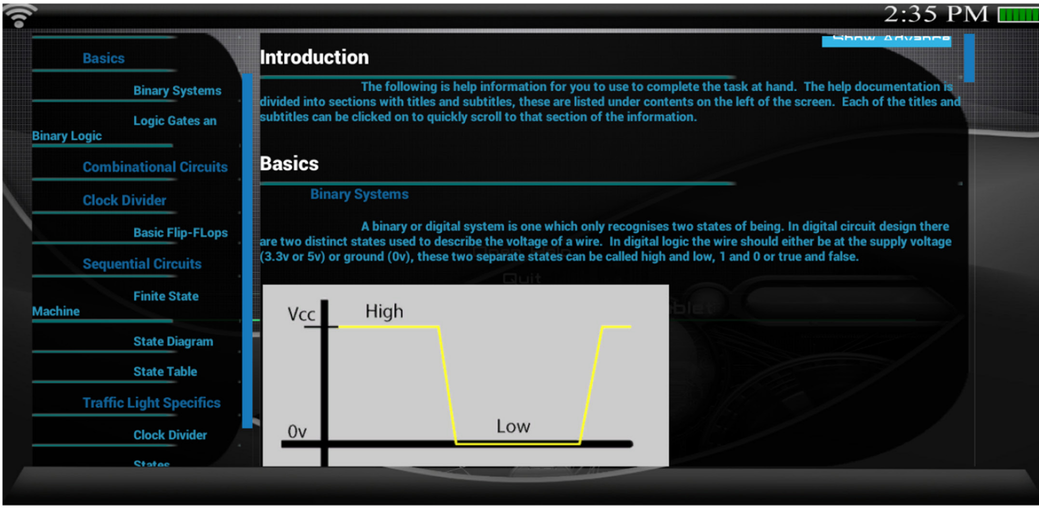

FIGURE 7 An excerpt from the help documentation in the initial game, before personalization and the addition of the AI support agent.

images (Figure 8b) or text with examples (Figure 8c) based on their past performance. This updated assistance is more focused on students' measured areas of difficulty, ultimately making the in-game help documentation significantly more useful for students.

Unexpected issues

To answer RQ2, we consider many unexpected issues that arose during implementation and testing and briefly detail them here to help future research potentially avoid or prepare for these issues. First and foremost, performance considerations are a major factor when dealing with virtual environments like those in our serious game. All students who participated in our study had personal computers, but many had lower-performing systems, which led to performance issues that impacted some students' experiences. Further, some students' computers had operating systems or hardware that we were unable to test on leading to compatibility issues that also impacted several students' experiences. With these considerations, future research on similar systems requires thorough testing on a variety of platforms, as well as a greater emphasis on performance optimization to ensure that students can participate properly.

Another significant issue was game distribution and Internet connectivity. To play the game, students first needed to download and install it. This process took a significant portion of the classroom time allocated for testing, giving students less time to interface with the game and complete their lab assignment. After realizing this issue during initial pilot testing, we began instructing students to download necessary software ahead of time to avoid taking up valuable classroom time.

Finally, no amount of development and testing can prepare such an educational system for the massive variety of interactions that students will attempt. To name only a few, students have: (1) given incorrect inputs to file submissions despite clear instructions; (2) discovered broken geometry in the virtual world and becoming unable to progress; (3) found a specific and untested sequence of inputs to lock a menu and prevent progression; and (4) had unusual Internet configuration on their personal computer leading to failed data collection through our servers. With such specific edge cases, it is impossible to test thoroughly

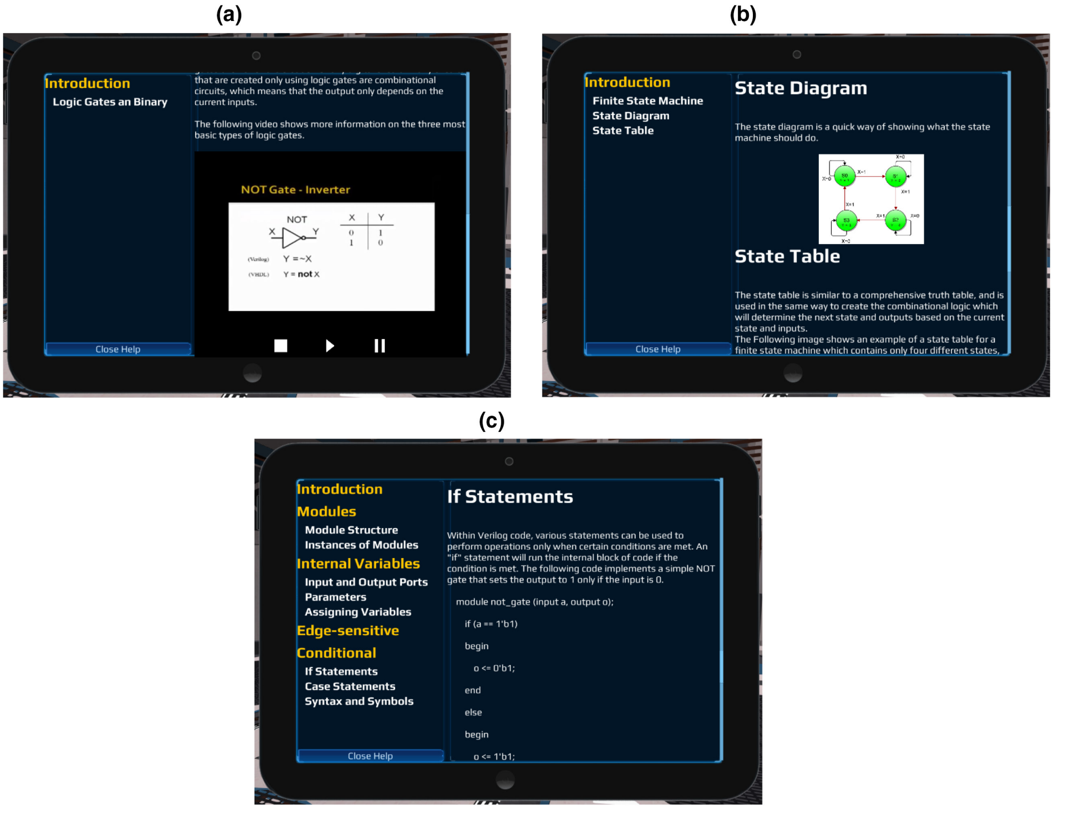

FIGURE 8 A few examples of the help documentation in the refined game showing a) video assistance; b) images; and c) example code. Based on AI decisions, this help is populated with different content for each student.

enough, and so as a final recommendation, it is very important to conduct observations and informal student interviews during testing to ensure that such issues can be fixed for future students.

METHODS

Our presentation here highlights 4 years of design-based research and iterative design changes based on student feedback. In each prior year, the game and system were tested in a first-year course for computer engineering students, with participating students completing feedback surveys after their experience. In spring of 2023, we also employed cognitive interviewing, otherwise known as the Think-Aloud protocol, to explore students' use of metacognitive strategies as they interacted with the game system (California State University, 2018; Knoll, 2018). The intent of the Think-Aloud protocol was to gain insight into students' cognitive and metacognitive processes as they played (Hacker, 2017), focusing specifically on how they interacted with the system and the ways in which it supported their learning. Throughout testing, our team also collected quantitative results on student performance, which are analysed in a separate publication (Hare, Tang & Ferguson, 2024).

Participants

For pilot testing, students in four lab sections of a first-year introductory university course had the option to participate in the study. Participating students signed consent forms and were randomly assigned to the standard lab assignment (control) or our current game with adaptive HCI system (treatment), while control group students did not interact with our software. As such, control group students are not represented in the results presented here, as their participation focused more on a qualitative comparison between the two groups as documented in prior literature (Hare, Tang & Ferguson, 2024). All students in the treatment group were invited to respond to feedback surveys. Table 1 shows participation numbers.

For cognitive interviewing, two students were selected from each of four lab sections to participate, and they met with the research team during their normal lab class meeting time. Interviews were held in a conference room down the hall from their lab session, with both students present at each meeting time. Students were informed beforehand and given a chance to ask questions and then asked to sign a consent form specific to the focus group activity.

Procedure

As stated, students in the treatment group were asked to complete surveys after their participation. Surveys asked about student interest in the game system and their perception of the system's usefulness. Additionally, members of the research team were present in classrooms during testing, directly observing students' interactions to find design issues.

Focus group interviews first provided students with a set of instructions as to how a think-aloud protocol works and an acknowledgement of the potential awkwardness of talking while you play a game. Examples were given for the students to understand the intention of the activity, and they were given the chance to ask questions and get used to the environment and research team before the activity began. While students played, the research team regularly prompted students to clarify their position in the game, their thoughts, the reasoning for their decisions or actions, and other related prompts. This approach focused on how students used the tools given, how they interacted with the system and the usefulness of the provided support.

Analysis

Student observation and feedback was key to the design process, directly informing the design iterations discussed in third section. Survey data helped to demonstrate shifts in students' responses as the system was adjusted year-by-year, and average survey responses from all years of testing are shown in Table 2. For Think-Aloud sessions, two devices were

TABLE 1 Number of student participants in our research study and optional survey over the past 4 years.

	Potential participants	Actual participants	Treatment group size	Survey responses
Year 1	117	39	19	16
Year 2	110	36	17	6
Year 3	52	49	28	15
Year 4	79	41	23	22

TABLE 2 Table of student survey responses by question, distributed among the possible choices listed.

Question	Response	Number of students			
		Year 1	Year 2	Year 3	Year 4
How interested were you in the game scenario?	Very	**9 (60%)**	2 (33%)	5 (33%)	2 (9%)
	Somewhat	6 (40%)	**3 (50%)**	**8 (53%)**	**16 (73%)**
	Not very	0 (0%)	1 (17%)	2 (13%)	4 (18%)
How interested were you in the design problem?	Very	**7 (47%)**	**4 (67%)**	3 (20%)	3 (14%)
	Somewhat	**7 (47%)**	1 (17%)	**10 (67%)**	**16 (73%)**
	Not very	1 (7%)	1 (17%)	1 (7%)	2 (13%)
How much did the problem seem like a realistic engineering task?	More than a textbook	**10 (77%)**	**4 (67%)**	**11 (73%)**	**11 (50%)**
	About the same as a textbook	3 (23%)	2 (33%)	4 (27%)	10 (45%)
	Less than a textbook	0 (0%)	0 (0%)	0 (0%)	1 (5%)
How interesting was the problem?	More than a textbook	**8 (62%)**	**3 (50%)**	7 (47%)	10 (45%)
	About the same as a textbook	5 (38%)	2 (33%)	**8 (53%)**	**12 (55%)**
	Less than a textbook	0 (0%)	1 (17%)	0 (0%)	0 (0%)
How much fun was the process?	More than a textbook	**13 (100%)**	**3 (50%)**	5 (33%)	**11 (50%)**
	About the same as a textbook	0 (0%)	2 (33%)	**9 (60%)**	**11 (50%)**
	Less than a textbook	0 (0%)	1 (17%)	1 (7%)	0 (0%)
To what extent did you have the resources to solve the problem?	More than a textbook	3 (23%)	**4 (67%)**	6 (40%)	9 (41%)
	About the same as a textbook	**10 (77%)**	2 (33%)	**8 (53%)**	**11 (50%)**
	Less than a textbook	0 (0%)	0 (0%)	1 (7%)	2 (9%)
How much did you learn?	More than a textbook	4 (31%)	**3 (50%)**	6 (40%)	7 (32%)
	About the same as a textbook	**8 (62%)**	2 (33%)	**9 (60%)**	**14 (64%)**
	Less than a textbook	1 (8%)	1 (17%)	0 (0%)	1 (5%)
To what extent did you increase your understanding of sequential circuits?	More than a textbook	4 (31%)	**2 (33%)**	**9 (60%)**	7 (32%)
	About the same as a textbook	**7 (54%)**	**2 (33%)**	6 (40%)	**15 (68%)**
	Less than a textbook	2 (15%)	**2 (33%)**	0 (0%)	0 (0%)

TABLE 2 (Continued)

		Number of students			
Question	**Response**	**Year 1**	**Year 2**	**Year 3**	**Year 4**
To what extent did you increase your interest in electrical engineering?	More than a textbook	**8 (62%)**	**3 (50%)**	**10 (67%)**	6 (27%)
	About the same as a textbook	4 (31%)	2 (33%)	5 (33%)	**16 (73%)**
	Less than a textbook	1 (8%)	1 (17%)	0 (0%)	0 (0%)

Note: Some students did not answer some questions, leading to inconsistent totals compared to those listed in Table 1. Percentages shown are based on the total number of students who answered the question in that year, and the highest value in each year is bolded.

used to record student and researcher comments, and transcripts were analysed for meaning making upon completion. Coding and analysis followed a constant comparative approach (Fram, 2013), using inductive coding and sorting to organize data around the primary attributes, and then comparing the developing concepts against the data iteratively until all major components were identified and explored throughout the data. The intention in analysis was to understand students' interactions with the game and their metacognitive or problem-solving strategies on display in the think-aloud.

RESULTS AND DISCUSSION

Central design ideas and system usability

As stated in the Method section, our presented results focus primarily on qualitative student feedback to support the insights presented in prior sections and to demonstrate positive responses to the changes made during our design process. With that in mind, we first focus on student feedback in the context of addressing RQ1; that is, feedback that supports the central design ideas presented in third section. For instance, after a test in year 2, one student's survey response requested "more direction … to let the user know what they should be doing," while another student mentioned that "the UI is poorly designed," causing some frustration when they participated in the study. After interface adjustments and added guidance aligning with our principles in third section, our open-ended surveys received no similar comments. Furthermore, Table 2 shows that students consistently rated the game as enjoyable throughout all 4 years of testing. The rating dropped slightly in year 3, but that coincides with the addition of the educational support and guidance, meaning that the educational aspects of the game were more noticeable, contributing to a lowered rating. After adding the gamified evaluations in year 4, the average score for enjoyment increased again.

Additional support for our design principles came from our focus group Think-Aloud protocol. As our study focuses on student–system interactions, codes in our focus group data specifically highlight students' problem-solving strategies and thought processes, as well as how they approach the game and use in-game tools. When students interacted with the provided AI assistance, it was generally helpful, with one student stating that they had previously been stuck on how to progress:

> But I felt like … once I did [find the help tools] it kind of gave me a little bit better understanding of, like, what we were supposed to do.
>
> (Focus Group 3 (FG-3))

Moving towards evaluation, more gamified evaluation methods do help in inducing a deeper learning process in students. One student, who had been stuck on the traffic light simulator in Figure 3b, shared a more developing approach in their play, noting:

> I felt like when I went to the help section I saw, like, the order of the traffic lights. So I was like oh oh, I guess that's what they want me to do. And then it kind of, it just kind of worked itself out, yeah.
>
> (FG-1)

Upon viewing a figure provided by the support system, the student grasped the correct steps. Another clear example of a student working through a problem in the game was seen in group two, where a student first voiced a question about the traffic light simulator and then answered their own question as they worked through it:

> Quick question. So like the outputs are like GN, RN or GE, which one is supposed to be like green light or like? Display when it's green was, like, is like, GN supposed to be like green, not…it's not or…Oh wait! Never mind. I got it, north and east. Yes. Yep, I remember that.
>
> (FG-2)

When students take time to explore and think through the game from a learning perspective, they do appear to develop a better understanding of the intentions of the game. And our observations and focus group interactions indicate that non-standard evaluation methods help to engage students in said exploration and thinking rather than presenting as a barrier to their gameplay. As an additional consideration from the focus group, it is clear that students in general do benefit from taking a short pause and thinking through or even discussing their learning process, and how best to encourage this metacognition and self-reflection exists as a potential focus for future design iterations.

Design issues

In addition to the positive insights gained from our Think-Aloud interviewing, the more personal format also helped to highlight further unexpected issues in our design. For instance, one student noted:

> I was really just trying to speed through the game, to be honest, so.
>
> (FG-3)

And based on observations, this is not a unique approach. Among similar students, standard evaluation measures such as quizzes tend to encourage students to throw themselves at the problem repeatedly while avoiding a longer learning process. At the same time, students with mentalities similar to the above tend to be experienced in virtual games, or as one student put it:

> "I'm a gamer, so…" with a quick shrug and a grin.
>
> (FG-2)

This specific mentality seems to place heavy focus on completion, and as such, students in this mindset are more inclined to skip documentation, miss instructions, and overall view evaluations and educational materials as barriers instead of aids to their learning process.

An important consideration then for gamified systems is how best to encourage these students to slow down and engage more with learning materials.

In addition to encouraging desired behaviour from students, interface design is also highlighted heavily in third section, ensuring that the game is usable and clear for all students, regardless of background. The necessity to do so was also highlighted in our focus group interviews. For instance, one student in the fourth focus group asked for more control options as the game started, noting that the introductory sequence of the game was overwhelming, requesting an option to pause the game:

> Like all my there was too much information passing by. I was like, whoa, why are [the cars] just stopping? So maybe if there was a button that's like, OK, stop, pause, simulate.
>
> (FG-1)

Or one student requesting more guidance in the game elements, sharing:

> I think clear directions and some of the activities would be beneficial.
>
> (FG-4)

The second remark indicates the necessity for not only clear guidance and instruction but also a good balance of where and how to present that information. For example, in some of the subgames, students are given “how to play” guidance at the beginning of the game. However, if they skim over this guidance or forget some portion of it, adding functionality to review the guidance even after starting the subgame is a key part of improving usability. And familiarity with games (and virtual software in general) also factors in to this design. To notify students of a recommended button, the game utilized exclamation marks as an “importance” icon, and one student was very honest in sharing:

> I'm just very oblivious when it comes to basic, like, game stuff like that … if I saw like the exclamation point, I would honestly start panicking there … I'm like, oh, God, what did I do wrong?
>
> (FG-2)

This comment is indicative of the importance of fully explaining iconography through keys or tutorials. With better explanation, inexperienced students can more easily navigate the environment and engage with provided systems. And above all, the aforementioned feedback supports the idea that educational software must be accessible to all levels of technological familiarity.

Student surveys and design process effectiveness

The final goal of this paper as per RQ3 is to verify the effectiveness of our iterative design process with regard to students' perceptions of system usability and utility. As stated, Table 2 shows student responses on several questions related to interest and usability of the system. Results year-to-year on student surveys overall indicated a positive response to the system compared to textbook learning overall during our iterative design changes. In general, student interest did not trend particularly up or down as our design changes were made, although there was a shift away from “very” and towards the “somewhat” option. This shift could be due to the expanded educational aspects of the game over later years. However, students' responses regarding how fun the process remained consistent year-over-year, meaning the expanded educational components did not make the game noticeably less enjoyable throughout our iterative changes. And as one student stated during our focus group:

> I think it is definitely more entertaining than just like a regular lab for sure.
>
> (FG-3)

Notably, our question regarding the game providing useful resources had a higher average response of "about the same as a textbook" in earlier years. This coincides with updates to the educational support, causing students in year 2 onwards to feel more strongly that the system and supportive AI components provided greater resources than textbook learning. This is further reflected in the question regarding students' understanding of the presented content, where we saw less students respond negatively when asked if the system helped increase their understanding of the material. Meanwhile, students consistently indicate that the game is comparable to a textbook as an educational tool, indicating that despite the gamified aspects and irregular content presentation, students perceive our system as useful and effective in their classroom learning.

CONCLUSIONS

In this work, we focused on the iterative design of an HCI-based educational system for a first-year university electrical engineering course. Through student surveys over 4 years, we discussed positive shifts in student perception and opinions of our HCI-based system as we iteratively improved the student experience. Our focus groups conducted using a cognitive interviewing or Think-Aloud protocol provided valuable insights into the game experience and the thought processes of students as they played. Altogether, these results give us a varied perspective beyond what can be gathered by looking at numerical outcomes and statistical analyses, which all provide valuable insights into designing meaningful and useful HCI informed by the human element.

Ultimately, this research offers valuable insights into the interplay between gameplay, thinking and problem solving. For future research, we highlight the importance of student feedback, as well as the importance of fostering a learning environment that encourages thoughtful problem solving. We also plan to continue iterating our design, focusing on encouraging better self-reflection and metacognition in students, as well as further improvements to the AI-based student support system. Finally, with our design nearing completion, future studies aim to explore the wider statistical impact of system implementation in classrooms, the long-term impacts on students' performance and overall success within their program and the effects of implementation with a wider and more diverse student population.

FUNDING INFORMATION

This research was supported by the National Science Foundation under grant number 1913809.

CONFLICT OF INTEREST STATEMENT

The authors declare no potential conflict of interests.

DATA AVAILABILITY STATEMENT

All data mentioned in this paper are available from the authors upon reasonable request.

ETHICS STATEMENT

All data used in this paper were collected with approval of the Rowan University IRB, #PRO-2021-315. Student data have been stripped of identifying information before analysis and publication. All students were made aware that participation was voluntary, and alternatives to the research study were offered to students who chose not to participate.

ORCID

Ryan Hare https://orcid.org/0000-0003-3318-2034

REFERENCES

Ahmad Faudzi, M., Che Cob, Z., Omar, R., Sharudin, S. A., & Ghazali, M. (2023). Investigating the user interface design frameworks of current mobile learning applications: A systematic review. *Education Sciences*, *13*(1), 94. https://doi.org/10.3390/educsci13010094

Alloghani, M., Al-Jumeily, D., Hussain, A., Aljaaf, A. J., Mustafina, J., & Petrov, E. (2018). Application of machine learning on student data for the appraisal of academic performance. In *2018 11th International Conference on Developments in eSystems Engineering (DeSE)* (pp. 157–162). Cambridge, UK.

Anantharaman, H., Mubarak, A., & Shobana, B. (2018). Modelling an adaptive e-learning system using LSTM and random forest classification. In *2018 IEEE Conference on e-Learning, e-Management and e-Services (IC3e)* (pp. 29–34). Langkawi, Malaysia.

Anderson, L., Krathwohl, D., Airasian, P., Cruikshank, K., Mayer, R., Pintrich, P., Raths, J., & Wittrock, M. (2000). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Pearson.

Andrade, H. L. (2019). A critical review of research on student self-assessment. *Frontiers in Education*, *4*, 468374. https://doi.org/10.3389/feduc.2019.00087

Barr, M. (2020). A cross-sectional study of video game play habits and graduate skills attainment. *Research in Learning Technology*, *28*. https://doi.org/10.25304/rlt.v28.2326

Bond, M., & Bedenlier, S. (2019). Facilitating student engagement through educational technology: Towards a conceptual framework. *Journal of Interactive Media in Education*, *1*(1), 1–14. https://doi.org/10.5334/jime.528

California State University. (2018). *Cognitive interview protocol template*. https://cair.org/wp-content/uploads/sites/474/2018/12/Kolbe-A-Framework-for-Evaluating-the-Success-of-a-Dashboard-Tool-Cog....pdf

Chaudhry, M. A., & Kazim, E. (2022). Artificial intelligence in education (AIEd): A high-level academic and industry note 2021. *AI and Ethics*, *2*(1), 157–165. https://doi.org/10.1007/s43681-021-00074-z

Cruz, Y. P., Collazos, C. A., & Granollers, T. (2015). The thin red line between usability and user experiences. In *Interacción '15: Proceedings of the XVI International Conference on Human Computer Interaction*. New York, NY.

Duvaud, S., Gabella, C., Lisacek, F., Stockinger, H., Ioannidis, V., & Durinx, C. (2021). Expasy, the Swiss bioinformatics resource portal, as designed by its users. *Nucleic Acids Research*, *49*(W1), W216–W227. https://doi.org/10.1093/nar/gkab225

Faghih, B., Azadehfar, M., & Katebi, S. (2013). User interface design for e-learning software. *International Journal of Soft Computing and Software Engineering*, *3*(3), 786–794. https://doi.org/10.48550/arXiv.1401.6365

Fram, S. M. (2013). The constant comparative analysis method outside of grounded theory. *NSUWorks*, *18*(1), 1–25. https://doi.org/10.46743/2160-3715/2013.1569

Franzwa, C., Tang, Y., Johnson, A., & Bielefeldt, T. (2014). Balancing fun and learning in a serious game design. *International Journal of Game-Based Learning*, *4*, 37–57. https://doi.org/10.4018/ijgbl.2014100103

González-González, C. S., Toledo-Delgado, P., & Muñoz-Cruz, V. (2015). Agile human centered methodologies to develop educational software. *DYNA*, *82*(193), 187–194. https://doi.org/10.15446/dyna.v82n193.53495

Hacker, D. (2017). The role of metacognition in learning via serious games. In R. Z. Zheng, & M. K. Gardner, (Eds.), *Handbook of research on serious games for educational Applications* (pp. 19–40). IGI Global.

Hare, R. & Tang, Y. (2023). Reinforcement learning with experience sharing for intelligent educational systems, In *IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 1431–1436), Honolulu, Oahu, HI, USA. https://doi.org/10.1109/SMC53992.2023.10394095

Hare, R., Tang, Y. & Ferguson, S. (2024). An intelligent serious game for digital logic education to enhance student learning. *IEEE Transactions on Education*, *67*(3), 387–394. https://doi.org/10.1109/TE.2024.3359001

Hasani, L. M., Sensuse, D. I., Kautsarina, & Suryono, R. R. (2020). User-centered design of e-learning user interfaces: A survey of the practices. In *2020 3rd International Conference on Computer and Informatics Engineering (IC2IE)* (pp. 1–7). Yogyakarta, Indonesia.

Kalmpourtzis, G., & Romero, M. (2020). Constructive alignment of learning mechanics and game mechanics in serious game design in higher education. *International Journal of Serious Games*, *7*(4), 75–88. https://doi.org/10.17083/ijsg.v7i4.361

King, D., & Ritchie, S. M. (2012). Learning science through real-world contexts. In K. G. Tobin, B. J. Fraser, & C. J. McRobbie (Eds.), *Second international handbook of science education* (pp. 69–79). Springer.

Knoll, T. (2018). The think-aloud protocol. In A. Drachen, P. Mirza-Babaei, & L. Nacke (Eds.), *Games User Research* (pp. 189–202). Oxford University Press. https://doi.org/10.1093/oso/9780198794844.003.0012

Kumar, D., Haque, A., Mishra, K., Islam, F., Mishra, B. K., & Ahmad, S. (2023). Exploring the transformative role of artificial intelligence and metaverse in education: A comprehensive review. *Metaverse Basic and Applied Research*, *2*, 55. https://doi.org/10.56294/mr202355

Limna, P., Jakwatanatham, S., Siripipattanakul, S., Kaewpuang, P., & Sriboonruang, P. (2022). A review of artificial intelligence (AI) in education during the digital era. *Advance Knowledge for Executives*, *1*, 1–9.

Lin, C. C., Huang, A. Y. Q., & Lu, O. H. T. (2023). Artificial intelligence in intelligent tutoring systems toward sustainable education: A systematic review. *Smart Learning Environments*, *10*(1), 1–22. https://doi.org/10.1186/s40561-023-00260-y

Miya, T. K., & Govender, I. (2022). UX/UI design of online learning platforms and their impact on learning: A review. *International Journal of Research in Business and Social Science (2147–4478)*, *11*(10), 316–327. https://doi.org/10.20525/ijrbs.v11i10.2236

Priyadarshini, P. (2024). The impact of user interface design on user engagement. *International Journal of Engineering Research & Technology*, *13*(3). https://doi.org/10.17577/IJERTV13IS030232

Puntambekar, S. (2022). Distributed scaffolding: Scaffolding students in classroom environments. *Educational Psychology Review*, *34*(1), 451–472. https://doi.org/10.1007/s10648-021-09636-3

Sáiz Manzanares, M. C., Rodríguez Arribas, S., Pardo Aguilar, C., & Queiruga Dios, M. Á. (2020). Effectiveness of self-regulation and serious games for learning stem knowledge in primary education. *Psicothema*, *32*(4), 516–524.

Stanton, J. D., Sebesta, A. J., & Dunlosky, J. (2021). Fostering metacognition to support student learning and performance. *CBE—Life Sciences Education*, *20*, fe3.

Yi, M., Wang, Y., Tian, X., & Xia, H. (2021). User experience of the mobile terminal customization system: The influence of Interface design and educational background on personalized customization. *Sensors*, *21*(7), 2428. https://doi.org/10.3390/s21072428

Zheng, L., Long, M., Zhong, L., & Gyasi, J. F. (2022). The effectiveness of technology-facilitated personalized learning on learning achievements and learning perceptions: A meta-analysis. *Education and Information Technologies*, *27*(8), 11807–11830. https://doi.org/10.1007/s10639-022-11092-7

How to cite this article: Hare, R., Ferguson, S., & Tang, Y. (2024). Enhancing student experience and learning with iterative design in an intelligent educational game. *British Journal of Educational Technology*, *00*, 1–18. https://doi.org/10.1111/bjet.13526