# An Intelligent Serious Game for Digital Logic Education to Enhance Student Learning

Ryan Hare, *Graduate Student Member, IEEE*, Ying Tang, *Senior Member, IEEE*, and Sarah Ferguson

***Abstract*—*Contribution:*** **A general-purpose model for integrating an intelligent tutoring system within a serious game for use in higher education. Additionally, this article also offers discussions of proper serious game design informed by in-classroom observations and student responses.**

***Background:*** **Personalized learning in higher education has become a key issue when working to improve student performance. By combining an intelligent tutoring system within a serious game, students can be engaged in their learning through gamified lessons while simultaneously receiving personalized and timely scaffolding to support their learning. Furthermore, related systems have not explored a general-purpose model for this type of system that can apply to any game or domain.**

***Intended Outcomes:*** **The combined intelligent tutoring system and serious game is well-received by students as determined by student surveys. Furthermore, students show better engagement in the given material and better performance on pre-post-intervention content tests.**

***Application Design:*** **The proposed system is a modular, general-purpose approach for integrating an intelligent tutoring system into any serious game for education. Using the machine learning paradigm of reinforcement learning, the system can adapt to student responses to improve future scaffolding.**

***Findings:*** **The results of the in-classroom testing are promising. Students who interacted with the intelligent game showed improved performance on content tests and positive responses on surveys regarding system usability and utility. This article also shows that students who used the intelligent game took less time and attempts to complete game sections, owing to the utility of the personalized support.**

***Index Terms*—Computer engineering, computer-based instruction, educational software, games, higher education.**

## I. Introduction

As EFFORTS in academia push toward education that is more widespread, accessible, and effective, one key issue is to tailor student learning experiences to individual needs, known as the personalization of student education [1]. This issue is especially prevalent in engineering education, where new concepts are often built heavily on prior classwork, and student difficulties with foundational knowledge can compound into poor understanding and classroom performance if not properly addressed [2], [3]. Nevertheless, the prevailing inclination in education is to adopt a one-size-fits-all approach that assumes all students learn in the same way. Unfortunately, this approach sacrifices educational effectiveness, as it fails to address the diverse range of potential issues and misunderstandings that students may encounter [4].

To focus on improving educational efficacy, it is necessary to explore approaches that offer more personalized tutoring to students who have outstanding or unexpected issues that the one-size-fits-all approach does not help with. However, personalized tutoring is time-consuming to plan, and requires significant effort on the part of the instructor to make sure that students are receiving appropriate support. One solution to this problem that has shown much potential in recent years is automated student tutoring using intelligent tutoring systems (ITSs) [5] and serious games (SGs) [6]. ITSs use data collection, student modeling, and computational intelligence to measure and predict a student's knowledge in a given subject and to offer personalized assistance that is tailored to that specific student. SGs, meanwhile, are defined as any game that serves a primary purpose other than entertainment. SGs offer the added benefit of gamification [7], [8], which can help to increase student engagement in and excitement about the subject matter, in turn leading to more knowledge retention.

This article discusses a domain-specific SG called Gridlock which integrates a problem-focused approach to educating students in digital logic design. Motivated by the above remarks about personalized education, SGs, and ITSs, Gridlock is further integrated with a reinforcement learning (RL) engine referred to as the personalized instruction and need-aware gaming (PING) system. Building on prior work [9], the current system is built to be more modular and useful to students with many usability improvements made to the Gridlock SG. Then, for any game integrated with the PING system, the RL agents can model student performance and adaptively provide timely assistance to students. With RL, the PING system can self-correct and self-learn how to properly interact with students simply by interacting with students. And while the discussions and results in this article focus on an implementation of the PING system within Gridlock, the system is built as a general-purpose, modular system that can work within any SG.

The main area of interest for this research is the student reception and educational merit of both Gridlock and the PING system. As such, this article describes both and focuses on

Manuscript received 30 May 2023; revised 13 December 2023; accepted 27 December 2023. Date of publication 28 February 2024; date of current version 4 June 2024. This work was supported in part by the National Science Foundation under Grant 1913809, and in part by the U.S. Department of Education Graduate Assistance in Areas of National Need (GAANN) under Grant P200A180055. *(Corresponding author: Ying Tang.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Rowan University IRB under Application No. Pro-2021-315.

Ryan Hare and Ying Tang are with the Department of Electrical and Computer Engineering, Rowan University, Glassboro, NJ 08028 USA (e-mail: tang@rowan.edu).

Sarah Ferguson is with the Department of Education, Rowan University, Glassboro, NJ 08028 USA.

Digital Object Identifier 10.1109/TE.2024.3359001

student performance and survey results from an in- classroom evaluation of both the PING system and Gridlock.

To that end, Section II provides a review of recent literature in this area. Section III discusses both Gridlock and the PING system in more detail regarding both the educational backing behind the design as well as changes made based on student feedback. Section IV provides a description of and results from pilot testing and in-classroom evaluations, followed by conclusions in Section V.

## II. Related Work

The use of both SGs (SGs) and ITSs for educational purposes has been an active research topic for many years [10]. When it comes to educational adaptation, the goal is simple; emulate a human tutor. For truly personalized learning, the system must be able to collect student data and make informed decisions on what tasks, feedback, help, or hints to provide to the student to ensure that the student's educational process is effective.

As SGs for education have been a common research topic, there are a number of SGs that take different approaches to education. Papadimitriou et al. [11] presented promising results from FuzzEG, a domain-specific game for programming that used fuzzy logic to modify quiz difficulty and game scenarios. Takahashi et al. [12], meanwhile, demonstrated an adaptive system for non-player character dialogue to help improve education. Hendrix et al. [13] provided a method for adaptive difficulty balancing for SGs to help maintain player engagement. Tan et al. [14] detailed a mobile app focused on small games for mathematics education. However, all of these systems provide implementations that are specific to their target domain or to their game. Meanwhile, the proposed work instead focuses on a general-purpose, modular system that can be applied to any SG and any educational domain.

In terms of more general-purpose approaches for adaptive and educational SGs, Dobrovsky et al. [15] presented a human-in-the-loop RL approach, where a human can provide corrective updates to the system to correct poor behavior. This approach, in fact, shares some similarities with the proposed method, but the proposed method attempts to avoid human-in-the-loop systems due to the necessary time pressure that those involved face when helping to improve system performance.

Another more general-purpose approach is presented by Bellotti et al. [16], who present an adaptive engine for SGs where an expert author is responsible for constructing and annotating tasks in the game. With their approach, the system automatically structures in-game missions to hit certain tasks and learning objectives. Compared to the proposed method, their approach is built for more open "sandbox" type games that have a much greater focus on exploration and nonlinear progression. Meanwhile, the proposed method focuses on more structured SGs where students progress in a more linear order.

## III. PING System Overview

Before going into detail of the PING system, it is important to discuss the key design aspects that were prioritized in relation to game structure and contents. Therefore, Section III-A discusses these pedagogical considerations; Section III-B provides a concise description of Gridlock, followed by its augmentation and adaptation in PING.

### A. Pedagogical Considerations in PING

The PING system discussed in this article is built to provide students with a real-world or otherwise contextual environment in which to solve a problem with the goal of improving engagement. Likewise, many other problems in STEM or wider education can be approached in a similar manner. As an initial step, it is important to identify what knowledge or skills the final system needs to impart to students. To that end, Gridlock applies a divide-and-conquer approach [17] where the goal is broken down into component parts that combine to address the overall subject matter. As an example, a SG may focus on educating students in a statistical analysis method. Students then would need to learn or prove their mastery of related topics, such as calculating averages, removing outliers, or other subskills, that contribute to the targeted pool of knowledge. Fig. 1 exemplifies this approach, showing how a SG and an ITS can combine, measure a student's performance, and provide timely feedback and support.

Another powerful learning tool that can be leveraged in these types of systems is repetition [18]. When students are tasked with similar tasks repeatedly, it helps to reinforce problem-solving steps, equations, or other relevant knowledge. Like in the classroom, practice is key. However, to maintain student engagement and ensure that proper learning occurs, it is helpful to add variation in repeated exercises. For example, in a quiz, varied questions or values from attempt to attempt helps to prevent the student from simply memorizing the answers. And even simple variation like varied backgrounds helps to add excitement to the student's experience.

In addition to divide-and-conquer and repetition, another key aspect of delivering lessons through SGs is enhancing student immersion or engagement in the subject matter. As mentioned, approaches that give students real-world tasks (or at least, realistic tasks) help to engage students and immerse them in a problem-solving situation [19]. At the same time, gamified elements are key to helping students stay engaged in the subject matter [20]. In other words, students should be so busy having fun or solving problems that they do not consider themselves in a "boring" educational setting. As such, it is important in any SG to balance educational content with gamified elements that prevent students from becoming distracted, bored, or otherwise disengaged in the educational content.

One final aspect of importance to both SGs and ITSs is data collection. To build systems that appropriately support student learning, data collection is key to capture students' content mastery, areas of difficulty, or mental state. With virtual environments, data collection is heavily enabled through easy access to student input data, movement through the environment, or actions with virtual objects and systems. Additionally, with the divide-and-conquer approach, it is easier for the system to measure content mastery in relation to small subsets of knowledge. In this way, a student can be given short

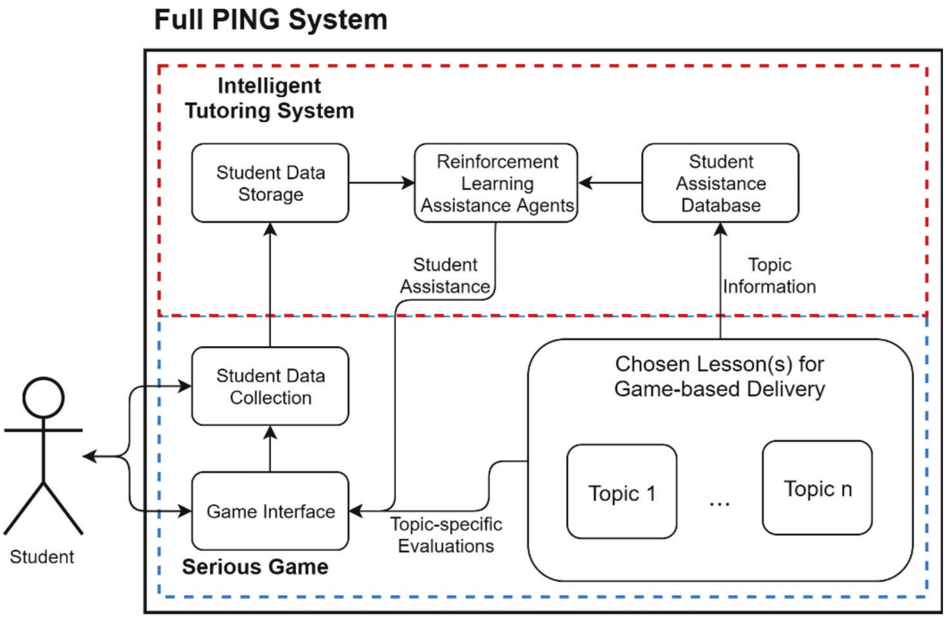


Fig. 1. Full diagram of the PING system showing the data flow between the ITS and the SG. Additionally shown is the divide-and-conquer approach to lesson planning in the SG.

Fig. 2. Traffic collision in Gridlock that helps students engage with the problem-solving process.

content tests on varying subjects to build an overall picture of their mastery of the target knowledge base.

### B. Gridlock

Gridlock is a domain-specific game developed for tutoring students on the fundamentals of digital logic design. Serving as an educational tool within the context of the Introduction to Digital Systems course, Gridlock specifically provides an alternative to the conventional lab where students design the logic controller for a traffic light. To help engage students in the narrative and add real-world context to the problem-solving process, students first witness a traffic collision caused by a failed traffic light, as shown in Fig. 2. This helps to motivate students to complete their task of redesigning the logic controller. Additionally, the virtual traffic light helps students feel more immersed in the problem, adding additional weight to the process while at the same time not putting adverse pressure on students.

Echoing the prior section, Gridlock uses a divide-and-conquer approach to teach students how to design the traffic light logic controller. Students instead learn and are tested on related topics, such as binary logic, logic circuit components, and syntax of the VerilogHDL programming language, which they use to create their final design. In this way, student knowledge can be tested individually on each topic to build an overall picture of the student's knowledge. Then, the student must prove mastery of all topics before progressing to the design stage. This structured approach also helps in implementing the ITS discussed in the next section. By recording separate performance variables for each topic, the system can more easily provide assistance in individual topics by leveraging the relevant student data.

Likewise, Gridlock also makes use of repetition, having students replay sections of the game until they demonstrate content mastery as indicated by the collected data. Some topics in Gridlock use simple quizzes to check student knowledge and help to reinforce concepts. As stated in the prior section, it is important that quiz questions are randomized between attempts, changing the question, possible answers, numerical values, and other aspects to ensure memorizing answers is not a valid strategy. Additionally, repetitive quizzes throughout every game section can be boring to students, especially when they expect a more gamified experience. To that end, Gridlock also replaces several quizzes with mini games that test the same knowledge. Figs. 3 and 4 show two of these mini-games, which are designed to test students on binary-to-decimal conversion and traffic light ordering, respectively. This way, the game still tests the same knowledge and still allows for repetitive exercises to help reinforce concepts. At the same time, the more gamified approach adds fun "break" periods between quizzes.

Gridlock is also a valuable tool for remote learning, and requires only a basic Internet connection to interface with


Fig. 3. Binary blaster, a mini game in Gridlock. In this game, students must match random binary numbers to their decimal counterparts.

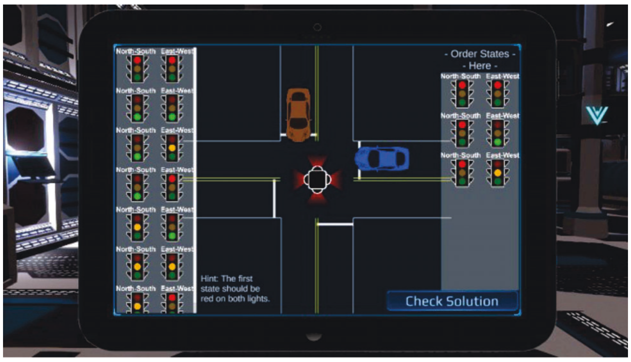


Fig. 4. Traffic light simulation mini game. In this game, students must properly order the lights in the traffic light to show they understand the design principles.


Fig. 5. In-game module selection interface. In this case, the student is recommended to do the binary blaster activity, and to get assistance on that section.

the PING system. When students enter, their game client is automatically assigned a unique ID to track their progress anonymously. The game client itself can run on most computer hardware, although there are some performance issues on older computers, such as low-frame-rates and long loading times. Whenever students are completing these activities or quizzes, the system can provide them with timely assistance as elaborated in the next section without instructor intervention. Furthermore, students can also request assistance at any time. In addition, the system can recommend specific activities and assistance, encouraging students to take certain paths through the game based on their performance. Fig. 5 shows how the system recommends certain sections to students, with small "notification" icons that indicate a recommended section, or a recommendation that the student request additional help.

### *C. PING System*

To provide timely assistance and section recommendations to students, Gridlock is integrated with the PING system. The PING system is a general-purpose framework for adaptive student assistance, applicable to any SG that utilizes a similar divide-and-conquer approach to problem solving. This way, the system has separate logic to provide assistance for each individual topic in the overall flow of the relevant game. For example, in Gridlock, the PING system has different logic for assisting students in binary-to-decimal conversion compared to providing assistance on VerilogHDL coding syntax.

To achieve adaptive student assistance, the PING system leverages RL. RL is an artificial intelligence method that uses trial-and-error to explore various approaches to a given problem and determine optimal behavior. In this case, the problem is defined as such: assist students and help them improve their performance. To that end, the RL agents first observe how a student is performing on content tests, mini games, or other assessments. Then, the agents select what assistance the student should receive. Finally, the agents are provided with feedback that shows if the decision was helpful or harmful to the student based on improvements in the student's measured performance. In the case of Gridlock, each topic has an RL agent that provides assistance within that topic. That way, each topic can have different logic, different priorities for what assistance to provide, and different data metrics that indicate better or worse-student performance.

Then, for the PING system, it is assumed that any game utilizing the system collects student data. In this work, the PING system was run on a remote server where each unique game client is assigned a random ID. Then, anonymous student data are sent to the server for the PING system to process, sending responses back to game clients. While the specific form of this data may vary, it is important that the data: 1) be relevant to and indicative of student performance in each topic and 2) be recorded for each topic in the game. For example, in Gridlock, each quiz or mini-game records several variables regarding student performance.

1) Total score on the quiz or mini-game pertaining to that section.
2) Total time taken to complete the quiz or mini game.
3) Emotional indicators while interacting with the content, measured by facial emotion recognition via webcam images using a recognition system built in to the game client.
4) Student inputs, including key presses and mouse movements, to indicate engagement with the system.
5) Index of quiz questions asked.

So, to summarize, the RL agents used in the PING system observe student performance as indicated by performance

TABLE I
EVALUATION PLAN MATRIX

| Research Questions | Evaluation Questions | Evaluation Measures |
|---|---|---|
| To what extent is the personalization tool and problem-solving content in the PING system useful to student learning | • Frequency of game tools being used by students<br>• Open-ended questions on why students like/dislike a particular tool | • Surveys of the utility and usability of the PING system |
| To what extent does the PING system with personalized, contextually and emotionally sensitive support in fostering student interests in engineering problem-solving | • The realism of games in delivering real-world engineering problems<br>• How fun and interesting does the problem-solving process in game compared to working out of a textbook/lab instruction | • Surveys of student interests in game learning |
| To what extent is the student learning improved by the PING system in general | • What do student reflections from student game experience reveal about their learning | • Game logs<br>• Pre-/post- content tests |

metrics gathered by the SG. Then, the agents provide students with appropriate assistance chosen from a pool of possible assistance decisions. Finally, improvements in student performance are provided to the system as feedback that helps improve future decisions. While a detailed explanation of RL is beyond the scope of this article, readers are referred to [21] for more details. Ultimately, this approach allows the system to have adaptive behavior that adjusts to trends in student responses while also requiring less effort on the developer's behalf to define the student assistance logic for each topic.

## IV. RESEARCH METHODOLOGY

A quantitative pre/post experimental research design was used to assess how the PING system impacted student interests and learning. The research evaluation plan was then guided by three research questions as detailed in Table I. The hypothesis in this study was that students provided with customized, context/emotion-sensitive support from the PING-integrated game will have increased content knowledge, faster completion times, less errors, fewer iterations, and more efficient problem-solving ability. Ultimately, these students are expected to exhibit increased interests and motivation to pursue high-level engineering design studies.

### A. Study Context

Data for this study were collected at one large public university in the Mid-Atlantic region of the United States. Students were recruited from two courses over the three years of the study: 1) Introduction to Digital Systems and 2) Computer Architecture. The experimental research design was set up to evaluate the incremental changes to the game across time, and all participating students completed IRB-approved consent forms before participation. If students chose to participate, they were randomly assigned to either a standard lab assignment with no game, an old version of Gridlock, or the new version of Gridlock with full personalized support. For students who did not sign consent forms, they participated in a standard lab assignment, and were not included in this study. In year one, the control group was a group of students engaging in the traditional lab for these courses using FPGA boards to simulate traffic lights. The treatment group for year one was the game without the PING system adaptations, and this year of data collection was used to train the system for implementation in year two. For year two and three, the treatment condition became the adapted version of the PING system game, with a comparison group of the original game without the PING addition to evaluate the incremental change of adding the PING functions. In total, this resulted in a control group sample of $n = 29$, a non-adapted game sample of $n = 37$, and an adapted game sample of $n = 43$.

### B. Data Collection

Several instruments were developed and used for evaluation. One sought to determine how students perceived the usefulness of the various scaffolding tools built into the PING system. Particularly, students were asked how helpful the tools were, how easy and how often they used them. The same instrument also measures student attitude toward the game scenarios and learning environment. Typically, students were asked to compare Gridlock to learning contents out of a textbook in terms of fun, interest, resource, and learning.

As stated in Section II, Gridlock features the design of sequential circuits using state diagram, state table and Verilog. To assess students' understanding of the content knowledge and problem-solving ability, the pre-/post- tests on vending machine design were given to both treatment and control groups. Specifically, students were asked to use state diagram/table to analyze and design the logic of a vending machine that meets the given specifications, and then implement it using Verilog. Students were given pre-intervention tests after covering relevant content in their course, before performing the related lab assignment. Post-intervention tests were administered as a part of regular course quizzes given a week after the lab assignment was completed.

## V. RESEARCH RESULTS

### A. Utility and Usability of Scaffolding Tools

The instrument for the utility and usability of PING was given to the treatment group using the adapted version of

TABLE II
ATTITUDE TOWARD THE GAME SCENARIO AND LEARNING

| Compared to working this problem out of a textbook (%) | | | | | | |
|---|---|---|---|---|---|---|
| | Realistic engineering task | Interest in problem | Fun of the process | Access to resources | Amount learned | Career interest |
| **Less** | 0.02 | 0.02 | 0.05 | 0.07 | 0.05 | 0.02 |
| **Same** | 0.37 | 0.51 | 0.51 | 0.49 | 0.53 | 0.54 |
| **More** | 0.61 | 0.47 | 0.44 | 0.44 | 0.42 | 0.44 |

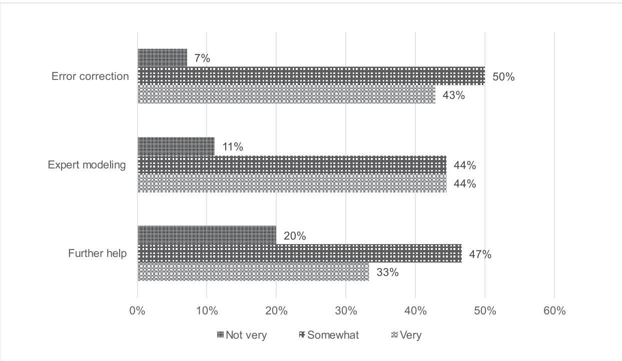


Fig. 6. Student opinions on the usefulness of the in-game tools, including error correction, expert modeling, and further help.

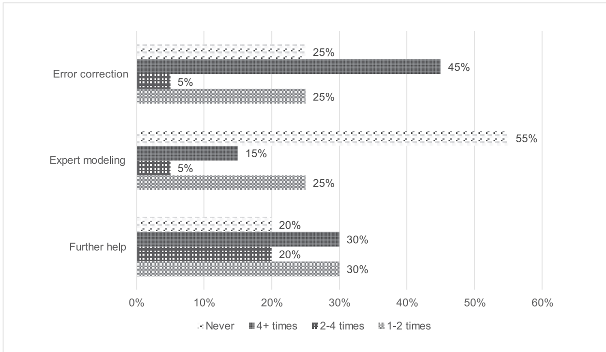


Fig. 8. Usefulness of in-game tools, as rated by the number of times students accessed those systems.

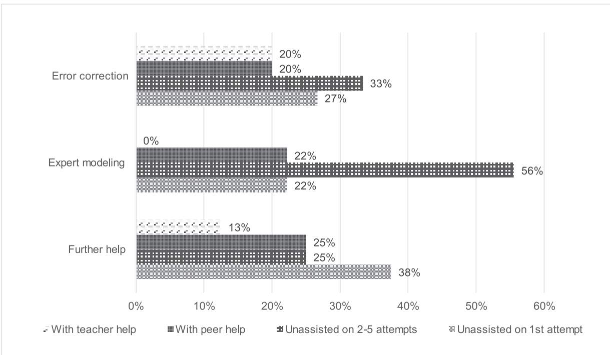


Fig. 7. Usability of the three in-game tools, as rated by how easy students found those systems to access.

the PING system only. Of the 43 students who participated, 20 students provided survey responses, which are shown in Figs. 6–8. As shown, the three tools were perceived useful in general with few problems figuring out how to use them. Of the three help tools, error diagnosis of student design was the most popular one (45% students used more than 4 times). Providing feedback on student performance, along with opportunities to repeat the "task-performance-feedback cycle" by allowing resubmission, is an effective practice in enhancing student self-regulation.

## *B. Interests*

Table II presents students' attitude toward the game scenario and learning environment. Overall, students consistently rated the game as comparable or better than working out of a textbook. And while learning contents were comparable to the ones in a textbook, a large percentage of students rated the game as more enjoyable and accessible compared to textbook learning, with very few students responding negatively in surveys. This aligns with the initial hypothesis that gamified education provides a better way to engage students. Furthermore, all surveys were administered anonymously, potentially reducing the bias toward positive responses from students aware of their participation in the research study.

## *C. Content Knowledge Understanding*

*1) Game Log Analysis:* The prior sections focused heavily on student perception, which may not necessarily algin with learning goals. To further verify the PING system, students' learning outcomes were analyzed through both game logs and pre-post-intervention content tests. First, student logs from their playthrough of the game were retrieved to compare the effects of the game system with and without the personalization. These comparisons focus on demonstrating the effect that iterative updates and design changes have had on students' game experiences. Outliers were removed with Grubb's test using an alpha of 0.05.

Initial comparisons capture the number of retries that a student made in each section of the game and are shown in Fig. 9. First, students in the adapted game completed the game with a significantly lower number of retries compared to these in the non-adapted version. The effect is especially prevalent among certain sections of the game, such as Section III, which deals with the mechanics of flip-flops, a key circuit element in Gridlock. The increased usability and accessibility of personalized help documentation within the game likely contributed to this outcome. By providing students with more

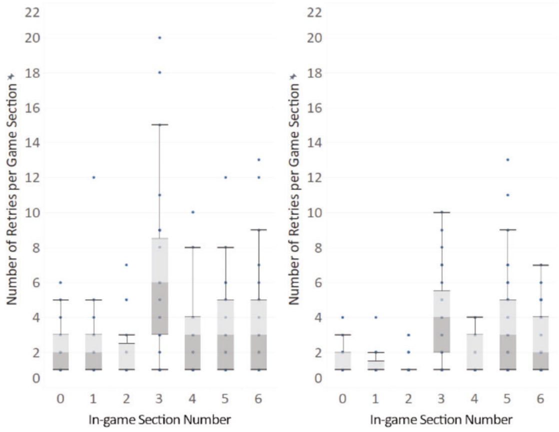


Fig. 9. Number of retries students made per section on the nonadapted game system (left) and the adapted game system (right).

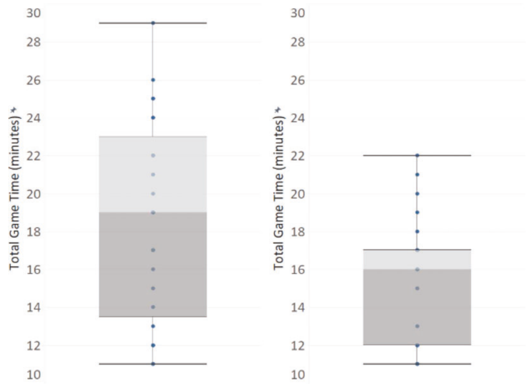


Fig. 10. Comparison of play time in minutes between the nonadapted game system (left) and the adapted system (right).

help documentation that is tailored closely to their individual needs, they were less inclined to require repeated attempts to complete game sections.

The second comparison focuses on time spent in-game by students as shown in Fig. 10. Clearly, students in the PING system spent less time moving through the educational content, owing to updates in the adaptive assistance, clearer instructions, and more accessible help documentation.

*2) Pre-/Post-Test Analysis:* Initially, group descriptive statistics were used to explore the differences in problem-solving scores by experimental groups over time. Table III presents that the adapted game treatment group students show positive change in their post-test scores for the engineering problem-solving measure over time with a large effect size (Cohen's $d$). The nonadapted game treatment group students also show positive change in their post-test scores when compared to pre-test though with a small effect size. Additionally, both game groups show higher-post-test scores when compared to the control group who did not participate in the game at all (Cohen's $d = 0.78$ non-adapted and $d = 1.15$ for adapted, both large effects), providing further support for both versions of the developed game.

To test the significance of this difference across all three groups, an ANCOVA was conducted to compare engineering

TABLE III
GROUP DIFFERENCE STATISTICS FOR ENGINEERING TASK PROBLEM-SOLVING MEASURE

| Group | Pre-Test Mean (*SD*) | Post-Test Mean (*SD*) | Cohen's *d* |
|---|---|---|---|
| Adapted Game ($n = 43$) | 4.72 (*3.04*) | 8.07 (*3.26*) | 1.06 |
| Non-Adapted Game ($n = 37$) | 6.04 (*2.54*) | 6.86 (*3.04*) | 0.30 |
| Control Group ($n = 29$) | 5.34 (*2.22*) | 4.59 (*2.76*) | -0.30 |

Note. Standard deviations in parentheses

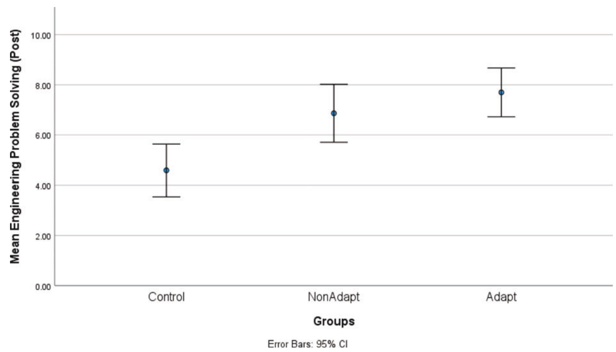


Fig. 11. Error bar graph of mean difference in problem-solving scores by group.

problem-solving post-test scores between the groups (control group, nonadapted game group, and adapted game group), using the pre-test scores as a covariate in the analysis. Cases with missing data at either pre-test or post-test were removed from the analysis, resulting in a usable sample size of 54 students for the ANCOVA. Assumption of normality for the continuous outcome variable and the covariate were checked using skewness and kurtosis values and were both found to be within range of $+/-$ 3.00. The homogeneity of variance assumption was checked using Levene's test and was supported, $F_{(2,51)} = 0.83$, $p = 0.44$.

There was a statistically significant difference in engineering problem-solving post-test scores between the three experimental groups, $F_{(3,50)} = 54.01$, $p = 0.02$. The effect size for this analysis was partial $\eta^2 = 0.17$, indicating that 17% of the variance in engineering problem-solving post-test scores can be explained by the experimental groups. Post hoc power analysis suggests the test was adequately powered ($1-\beta = 0.85$) with a medium sample size ($n = 54$). Based on this result, the full model is statistically significant with a medium practice significance. Fig. 11 provides a visual representation of these results with 95% confidence interval error bars, clearly showing the significant difference between the control group and the two treatment groups, with the strongest effect between the control and the adapted game group. Differences between the non-adapted game and the adapted game groups are not statistically significant in this analysis, as represented by the overlapping error bars.

Finally, the present study has a few limitations that may affect the extensibility of these findings. First, the number of students participating in the study could be considered a limiting factor, as it is possible that the conclusions presented are not extensible to the student population as a whole. However, these conclusions are significant based on the statistical tests presented, though more student participants would serve to bolster these conclusions. Furthermore, participants were limited to a specific geographical region, which may have impacted study results as students in different geographical regions can have different preferences, motivations, or interests. These differences would be even more notable when considering cultural differences across different countries. As such, the presented results may not extend fully into different geographical regions given varied student preferences. Finally, the presented work focused solely on quantitative results from the study. While these results do show statistical merit, a qualitative analysis on the collected data would further serve to improve the merit of this study, mainly because the effectiveness of education can be heavily dependent on student emotions and feelings. Further exploring the qualitative side of things with student thoughts, emotions, and opinions would improve the impact of this study.

## VI. Conclusion

This article overviews the PING system, an ITS specifically designed for seamless integration within an educational SG. The pedagogical insights applied in Gridlock, including divide-and-conquer approaches to problem solving, repetitive lessons with various delivery formats, and gamified student evaluation, are elaborated in detail. These strategies successfully strike a balance between learning and enjoyment within the game environment, as supported by student surveys, pre-/post-intervention assessment, and game log data. The findings in this article demonstrate that the PING system is user-friendly and useful with greater benefits compared to working on the same problem with paper and pencil. The purpose of the comprehensive student assessment in Gridlock is twofold: 1) it supplies vital data for PING to analyze student learning behaviors and 2) personalizes instructions to meet student individual needs. Second, the assessment process fosters student self-reflection, deepening their metacognitive awareness of their own learning. Thus, additional effort is needed to determine the metacognitive benefits of both SGs and the PING system [22].

For future work, as mentioned, this article focused entirely on quantitative results from the study. Qualitative analysis of the implemented PING system and of Gridlock is ongoing [23]. Future work also includes continuous enhancement and generalization of the PING system to suit a wider range of SGs and educational contexts [24]. This generalization would serve to benefit the research community, as the proposed system would better fit a wider range of domains or subjects.

## References

[1] Y. Tang, J. Liang, R. Hare, and F.-Y. Wang, "A personalized learning system for parallel intelligent education," *IEEE Trans. Comput. Soc. Syst.*, vol. 7, no. 2, pp. 352–361, Apr. 2020, doi: 10.1109/tcss.2020.2965198.

[2] L. Tetzlaff, F. Schmiedek, and G. Brod, "Developing personalized education: A dynamic framework," *Educ. Psychol. Rev.*, vol. 33, no. 3, pp. 863–882, 2020, doi: 10.1007/s10648-020-09570-w.

[3] S. A. Hart, "Precision education initiative: Moving towards personalized education," *Mind, Brain, Educ.*, vol. 10, no. 4, pp. 209–211, 2016, doi: 10.1111/mbe.12109.

[4] S. Goodman, and H. Bohanon, "A framework for supporting all students: one-size-fits-all no longer works in schools," *Amer. School Board J.*, pp. 1–4, Feb. 2018.

[5] A. C. Graesser, X. Hu, and R. Sottilare, "Intelligent tutoring systems," in *International Handbook of the Learning Sciences*. Abingdon, U.K.: Routledge, 2018, pp. 246–255, doi: 10.4324/9781315617572-24.

[6] Y. Zhonggen, "A meta-analysis of use of serious games in education over a decade," *Int. J. Comput. Games Technol.*, vol. 2019, pp. 1–8, Feb. 2019, doi: 10.1155/2019/4797032.

[7] M. Milosz and E. Milosz, "Gamification in engineering education—A preliminary literature review," in *Proc. IEEE Glob. Eng. Educ. Conf. (EDUCON)*, 2020, pp. 1975–1979, doi: 10.1109/educon45650.2020.9125108.

[8] F. F.-H. Nah, Q. Zeng, V. R. Telaprolu, A. P. Ayyappa, and B. Eschenbrenner, "Gamification of education: A review of literature," in *Proc. 1st Int. Conf. HCI Bus.*, 2014, pp. 401–409, doi: 10.1007/978-3-319-07293-7_39.

[9] Y. Tang and R. Hare, "Evaluation of a game-based personalized learning system," in *Proc. ASEE Annu. Conf.*, 2021, pp. 1–11.

[10] R. Hare and Y. Tang, "Player Modeling and adaptation methods within adaptive serious games," *IEEE Trans. Comput. Soc. Syst.*, vol. 10, no. 4, pp. 1939–1950, Aug. 2023, doi: 10.1109/TCSS.2022.3203926.

[11] S. Papadimitriou, K. Chrysafiadi, and M. Virvou, "FuzzEG: Fuzzy logic for adaptive scenarios in an educational adventure game," *Multimedia Tool Appl.*, vol. 78, no. 22, pp. 32023–32053, 2019, doi: 10.1007/s11042-019-07955-w.

[12] T. Takahashi, K. Tanaka, and N. Oka, "Adaptive mixed-initiative dialog motivates a game player to talk with an NPC," in *Proc. 6th Int. Conf. Hum.-Agent Interact.*, 2018, pp. 153–160, doi: 10.1145/3284432.3284436.

[13] M. Hendrix, T. Bellamy-Wood, S. McKay, V. Bloom, and I. Dunwell, "Implementing adaptive game difficulty balancing in serious games," *IEEE Trans. Games*, vol. 11, no. 4, pp. 320–327, Dec. 2019, doi: 10.1109/tg.2018.2791019.

[14] C. W. Tan, L. Ling, P.-D. Yu, C. N. Hang, and M. F. Wong, "Mathematics gamification in mobile app software for personalized learning at scale," in *Proc. IEEE Integr. STEM Educ. Conf. (ISEC)*, 2020, pp. 1–5, doi: 10.1109/isec49744.2020.9397846.

[15] A. Dobrovsky, U. M. Borghoff, and M. Hofmann, "An approach to interactive deep reinforcement learning for serious games," in *Proc. 7th IEEE Int. Conf. Cogn. Infocommun. (CogInfoCom)*, 2016, pp. 000085–000090, doi: 10.1109/coginfocom.2016.7804530.

[16] F. Bellotti, R. Berta, A. De Gloria, and L. Primavera, "Adaptive experience engine for serious games," *IEEE Trans. Comput. Intell. AI Games*, vol. 1, no. 4, pp. 264–280, Dec. 2009, doi: 10.1109/tciaig.2009.2035923.

[17] D. Delgado-Gómez, F. González-Landero, C. Montes-Botella, A. Sujar, S. Bayona, and L. Martino, "Improving the teaching of hypothesis testing using a divide-and-conquer strategy and content exposure control in a gamified environment," *Mathematics*, vol. 8, no. 12, p. 2244, 2020, doi: 10.3390/math8122244.

[18] S. H. Kang, "Spaced repetition promotes efficient and effective learning," *Policy Insights From Behav. Brain Sci.*, vol. 3, no. 1, pp. 12–19, 2016, doi: 10.1177/2372732215624708.

[19] E. H. J. Yew and K. Goh, "Problem-based learning: An overview of its process and impact on learning," *Health Prof. Educ.*, vol. 2, no. 2, pp. 75–79, 2016, doi: 10.1016/j.hpe.2016.01.004.

[20] R. Smiderle, S. J. Rigo, L. B. Marques, J. A. Peçanha De Miranda Coelho, and P. A. Jaques, "The impact of gamification on students' learning, engagement and behavior based on their personality traits," *Smart Learn. Environ.*, vol. 7, no. 1, p. 3, 2020, doi: 10.1186/s40561-019-0098-x.

[21] R. Hare and Y. Tang, "Petri Nets and hierarchical reinforcement learning for personalized student assistance in serious games," in *Proc. Int. Conf. Cyber-Phys. Soc. Intell. (ICCSI)*, 2022, pp. 733–737, doi: 10.1109/iccsi55536.2022.9970680.

[22] Y. Tang and R. Hare, "Combining gamification and intelligent tutoring systems in a serious game for engineering education," 2023, *arXiv:2305.16568*.

[23] R. Hare, Y. Tang, and C. Zhu, "Combining gamification and intelligent tutoring systems for engineering education," in *Proc. IEEE Frontiers Educ. Conf. (FIE)*, Oct. 2023, pp. 1–5, doi: 10.1109/fie58773.2023.10343378.

[24] R. Hare and Y. Tang, "Hierarchical deep reinforcement learning with experience sharing for metaverse in education," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 53, no. 4, pp. 2047–2055, Apr. 2023.