Singapore Computer Society
International Journal of Information Technology
Volume 31, Issue 1, 2025

Leveraging Multimodal Large Language Models for Gamified Learning in Higher Education: A Case Study in Computer Science

Sixuan Du[1], Yang Qiu[2], Zhiqi Shen[2] and Siyuan Liu[2]

50 Nanyang Avenue,
Nanyang Technological University, Singapore

[1]sixuandu96@outlook.com,
[2]{qiuyang, zqshen, syliu}@ntu.edu.sg

Abstract

Gamification has gained attention in education for their ability to enhance student engagement through interactive elements such as storytelling, challenges, and rewards. However, their widespread adoption is often hindered by the resource-intensive nature of development, which requires expertise spanning pedagogy, game design, and technical implementation. Recent advances in multimodal large language models (MLLMs) offer a promising avenue to address these challenges by enabling collaborative support across textual, visual, and code-based game artifacts, thereby reducing the time and effort required for serious game creation. In this work, we explore the use of MLLMs to facilitate the design and implementation of EscapeDA, an escape-room-style serious game prototype developed to promote pre-tutorial engagement in a computer science module, Data Structures and Algorithms. Through multimodal interactions, the MLLM supports narrative construction, challenge design, and implementation-level reasoning, assisting developers in iteratively refining game mechanics and content. By gamifying tutorial materials, EscapeDA encourages meaningful student interaction with learning resources, addressing low preparation and participation commonly observed in flipped classroom settings. A pilot study demonstrates that

MLLM-facilitated game development can effectively lower design barriers while supporting engaging serious game experiences, highlighting the potential of multimodal LLMs to scale gamification efforts in higher education.

Keyword: MLLMs, Gamification, Education.

I. Introduction

In recent years, gamification has gained significant traction in educational settings due to its ability to enhance student engagement [18, 19, 25]. Among various gamification approaches, serious games represent a prominent category that embeds educational objectives within game-based experiences. By integrating instructional content with game elements such as storytelling, challenges, and rewards, serious games transform traditional learning activities into interactive and enjoyable experiences [3, 20]. Prior research has consistently identified student engagement as a critical factor in academic success, and serious games offer an effective means of sustaining learner motivation across diverse educational contexts [4]. Despite these advantages, the widespread adoption of serious games in education remains limited, largely due to the substantial time, effort, and expertise required for their design and development [33].

Designing an effective educational serious game is inherently complex and resource-intensive, requiring expertise spanning pedagogy, game design, and technical implementation. Educators and designers must carefully balance engaging gameplay with meaningful learning objectives through iterative processes involving narrative construction, puzzle design, interface layout, and system implementation. These high entry barriers discourage many educators from adopting serious games, even in disciplines where sustaining student engagement is particularly challenging [24, 29]. In higher education computer science programmes, for example, abstract and technically demanding topics such as data structures and algorithms often pose difficulties in maintaining student interest and encouraging active participation.

Recent advances in large language models (LLMs), such as the GPT series, offer a promising avenue to alleviate these development challenges. LLMs have demonstrated strong capabilities in creative content generation, scenario construction, and interactive collaboration, enabling designers to rapidly prototype narratives, puzzles, and instructional elements. Compared with traditional computer-aided design tools, LLMs support flexible, conversational interaction with human designers, potentially reducing development effort and accelerating iteration cycles. However, most existing applications of LLMs in educational game development rely primarily on text-based interaction, which limits their effectiveness in supporting serious game creation [10, 17]. Serious games inherently involve multimodal artifacts, such as visual layouts, interface elements, and implementation code, and ensuring alignment among these artifacts, gameplay mechanics, and educational objectives remains a non-trivial challenge.

In this work, we investigate how multimodal large language models (MLLMs) can be leveraged to facilitate the design and implementation of serious games in educational contexts. Specifically, we employ GPT-4V (GPT-4 with vision), a multimodal LLM capable of jointly processing textual and visual inputs, enabling integrated reasoning over narrative descriptions, game screenshots, and implementation-level representations. By supporting multimodal interaction, the model provides enhanced assistance throughout the serious game development pipeline, from narrative and puzzle design to interface-aware and logic-consistent refinement.

We adopt a higher-education computer science module, Data Structures and Algorithms, as a case study. This module is foundational in most computer science curricula and serves as a prerequisite for advanced modules such as Operating Systems, Databases, and Artificial Intelligence, making sustained student engagement particularly important for long-term academic success. The module is typically delivered through large-scale lectures complemented by smaller tutorial sessions (e.g., 20–30 students per group) aimed at reinforcing conceptual understanding. A flipped classroom approach is commonly employed, where students are expected to attempt tutorial questions prior to attending sessions [2]. In practice, however, many students attend tutorials without sufficient preparation, resulting in reduced participation and suboptimal learning outcomes.

To address this issue, we designed and developed EscapeDA, an escape-room-style serious game prototype, leveraging GPT-4V to facilitate its design and implementation. During development, the model was provided with both textual descriptions of learning objectives and game mechanics, as well as visual artifacts such as screenshots of game scenes and user interfaces. EscapeDA embeds tutorial content into interactive challenges, encouraging students to engage with learning materials prior to tutorial sessions in a more engaging and structured manner. We further conducted a pilot user study to evaluate the game's design and its impact on student engagement. Our findings indicate that GPT-4V–facilitated serious game development can substantially reduce design barriers while fostering strong student involvement, highlighting the practical potential of multimodal LLMs to support scalable serious game creation in higher education under constrained resources.

II. Related Work

Serious games, which combine educational objectives with engaging gameplay, have gained significant attention in recent years as a transformative approach to education [18, 19, 25]. These games integrate educational content with interactive storytelling, challenges, and rewards, aiming to make the learning process more engaging and effective. Research has demonstrated that serious games can enhance motivation [26], engagement [27], sense of achievement [11], knowledge retention [25], and academic performance [22]. However, designing an effective educational game involves a multidisciplinary approach, combining expertise in pedagogy, game design, and software development [33]. This complexity often makes the development process resource-intensive and inaccessible for many educators and institutions, especially those with limited budgets or technical capabilities. Another core challenge in serious game design is balancing educational objectives with engaging gameplay [5]. Overemphasizing educational content can make the game feel like a traditional learning task, reducing its appeal to students. On the other hand, focusing too much on entertainment risks detracting from the educational value, resulting in a game that is fun to play but ineffective as a learning tool [15]. Striking the right balance is critical to ensure that the game captivates learners while still meeting its pedagogical goals.

Large language models (LLMs) have emerged as powerful tools to address some of these challenges by streamlining game design and development processes. LLMs have shown potential in multiple aspects of game design, including enhancing creativity [1], generating game levels [30], constructing scenarios [14], and supporting game testing [23, 28]. For example, Värtinen et al. demonstrated that a fine-tuned GPT-2 model could autonomously generate task descriptions for video games with minimal human supervision [32]. Similarly, van Stegeren and Myśliwiec showed that GPT-2 could replicate the structural and stylistic language of “World of Warcraft” tasks, illustrating LLMs’ ability to produce domain-specific content [31]. By reducing the cognitive load on designers, LLMs enable greater focus on higher-level creative tasks, thereby accelerating the game development process.

More recently, research has shifted to iterative interactions with LLMs to enable real-time and dynamic game design. Lanzi et al. proposed a collaborative framework that integrates interactive evolution with ChatGPT to simulate human-centered design processes, receiving positive feedback from game designers [14]. S7milarly, Fulcini et al. investigated ChatGPT’s ability to devise gamification strategies in a software engineering context. Their findings highlighted LLM’s utility in providing feedback and identifying design risks, but they also emphasized that the model still requires significant human supervision to produce valid and complete solutions [8].

In educational contexts, LLMs have shown promise in supporting the development of serious games by facilitating content generation and early-stage design exploration. Nevertheless, ensuring alignment with learning objectives while maintaining learner engagement remains a significant challenge [13]. Fotaris et al. investigated the use of LLMs within a design-thinking framework for conceptualising serious games, but their study did not extend to the design or evaluation of a fully implemented educational game [7]. Moreover, existing work has not examined how multimodal LLMs, capable of reasoning over visual artefacts such as design diagrams and UI screenshots, can support the end-to-end design and development of serious games.

Building on this gap, our work leverages multimodal LLMs (MLLMs), e.g., GPT-4 with vision capabilities (GPT-4V) to design and develop an educational escape-room-style serious game prototype for a computer science module on Data Structures and Algorithms. By enabling the model to jointly reason over textual specifications and visual artefacts, such as level diagrams and interface screenshots, our approach supports iterative game design and implementation, with the aim of enhancing student engagement in pre-tutorial preparation.

III. Gamifying Educational Content with MLLMs

Game Design in Education Environment with MLLMs

As highlighted in [12, 21], design is generally a collaborative process aimed at solving problems, beginning with a design brief that outlines the overarching goals. The process unfolds in three phases: the divergent phase, where a range of potential ideas are generated to explore various possibilities; the emergent phase, where these ideas are iteratively refined, modified, and combined to produce innovative solutions; and the convergent phase, where the refined ideas are critically evaluated and finalized into a cohesive outcome. In educational settings, the main challenge in game design lies in ensuring alignment with the intended learning objectives. To address this, close coordination among course instructors, gamification designers, and developers is essential throughout the design process, as shown in Figure 1.

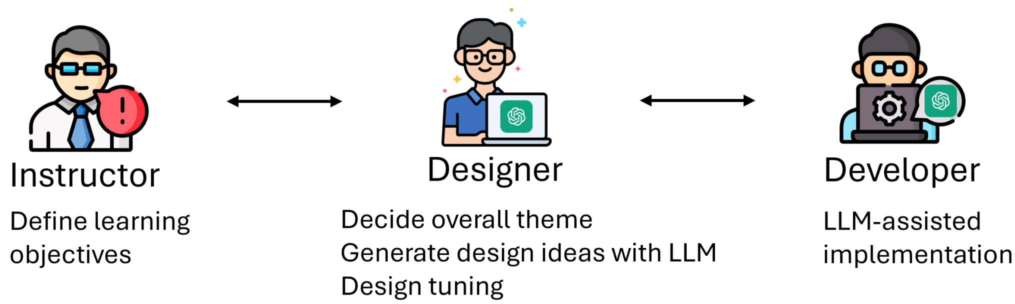


Figure 1. Proposed approach with roles and tasks

In the divergent phase, the instructor, who is well-versed in the educational content, begins by listing the specific learning objectives. This step ensures that the gamified elements align with the core learning goals. For example, if the objective is to help students understand an algorithm's logic, the design should focus on fostering conceptual reasoning rather than overwhelming students with coding tasks. These objectives are then shared with the gamification designer, who selects a theme from an existing art asset library and crafts detailed prompts for MLLMs to generate initial design ideas. The prompts are composed of four components: background information about the subject, learning content, associated learning objectives, and screenshots of the target game environment. These detailed inputs enable LLMs to provide tailored and contextually relevant suggestions to be aligned with the learning objectives, such as game mechanics, interactive challenges, and storylines that serve as a foundation for the game's development.

In the emergent and convergent phrases, the instructor, designer, and developer collaborate to refine and iterate on the MLLM-generated solutions. Leveraging on MLLMs, the solutions will be adapted to the educational context and resource constraints. For instance, if an initial idea involves complex gameplay mechanics that may exceed development resources, MLLMs will be prompted to suggest simpler alternatives that retain the intended educational value.

Case Study: The Design of EscapeDA

We applied this approach in designing a tutorial for the module "Data Structures and Algorithms". The specific tutorial was designed to deepen students' understanding of two core topics: the Depth-First Search (DFS) algorithm for graph traversal and Dynamic Programming strategy. Each topic was paired with a practice question, detailed in Appendix A. For the DFS question, the learning objective in the tutorial is to enhance students' comprehension of the algorithm's logic by asking them to produce the sequence of nodes visited for a given graph according to the DFS algorithm. For Dynamic Programming, the objective is to guide students in applying the strategy of breaking down complex problems into smaller subproblems.

The game design and development team consisted of the module instructor, a gamification researcher, and an undergraduate student hired as a part-time developer. GPT-4V was employed to facilitate design and development. To boost student motivation and engagement, the team avoided "shallow" gamification strategies, such as merely awarding points for correct answers. Rather than relying on GPT-4V to generate the design from scratch, during the initial divergent phase, the team decided on an escape room theme. This theme is particularly effective for gamifying tutorial questions, as it enables each puzzle to be closely aligned with a specific tutorial question. Furthermore, the escape room setup engaged students by motivating them to solve puzzles to progress and earn rewards, making the game both enjoyable and goal-oriented [6]. To further streamline the design process and narrow the scope of LLM-generated solutions, the team opted for a 2D environment. This environment was chosen for its rich asset library and flexibility, enabling efficient implementation.

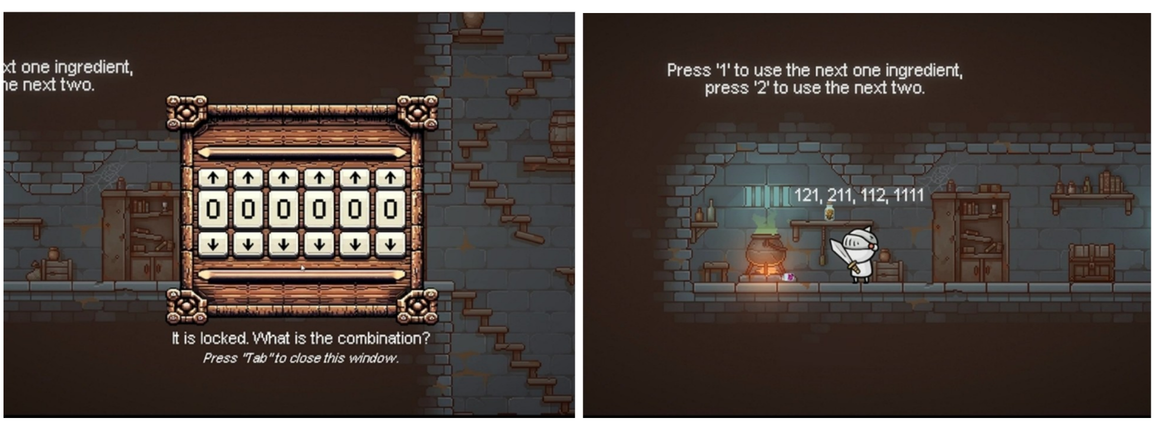
With the theme and environment chosen, prompts were crafted and passed to GPT-4V to generate initial game design ideas. These prompts incorporated background information about the tutorial, details of the tutorial questions (i.e., learning content), associated learning objectives, and screenshots of the target game environment. An example of such a prompt is provided in Appendix B. GPT-4V generated relevant and tailored design suggestions based on these inputs. In the emergent and convergent phases, the design was iteratively refined, and implementations were expedited using GPT-generated code snippets.

Facilitated by GPT-4V, we developed the educational game, EscapeDA, which is available as a playable online demo[1]. Set in a mysterious dungeon, players assume the role of an explorer navigating through puzzles and obstacles in an effort to escape. Following quests outlined in an in-game notebook, players complete tasks to advance through each puzzle. In the first puzzle as shown in Figure 2, players light candles in a specific sequence, simulating the Depth-First Search (DFS) algorithm. Starting from an initial node, players light candles one by one as they traverse deeper along a path. When they encounter a dead end, they backtrack and choose an unexplored path.

[1] https://annoymaous.itch.io/dfs-and-dynamic-programming-tutorial

Figure 2. Puzzle 1: Top-left shows the movement instructions in game Top-right shows the player discovering a chest. Bottom-left is the ‘Dungeon Survival Guide’ which outlines the task of lighting the candles in order. Bottom-right provides feedback when the candle-lighting order is incorrect.

The second puzzle focuses on Dynamic Programming as shown in Figure 3, set within a narrative involving potion mixing. Players are introduced to a puzzle where the number of potions corresponds to the Fibonacci sequence. They are tasked with implementing a program to calculate a specified Fibonacci number.



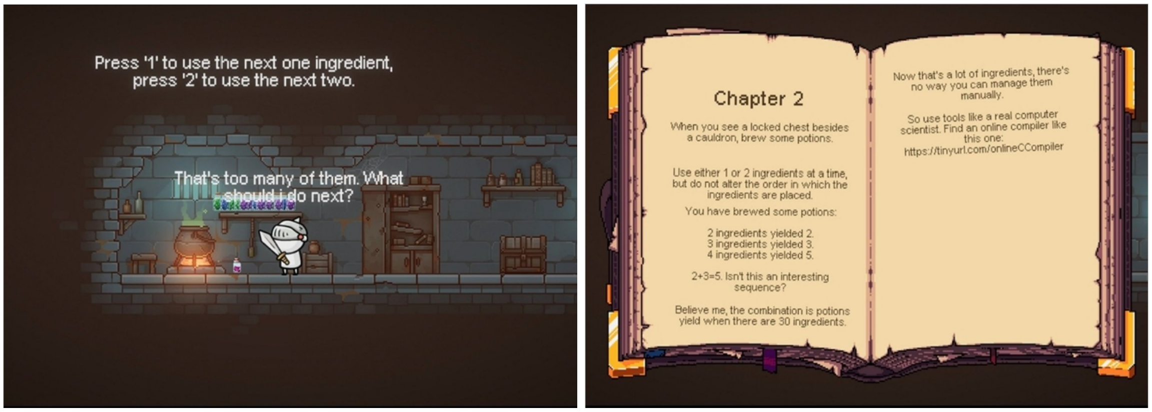


Figure 3. Puzzle 2: Top-left shows the locked chest, prompting the player to input the correct combination to unlock it. Top-right displays the potion-mixing interface, where the player inputs potion combinations using the keyboard. Bottom-left alerts the player that the number of potions entered exceeds the expected amount, asking what to do next. Bottom-right shows the quest book which outlines the potion combination problem, introduces the Fibonacci sequence concept, and presents the challenge of determining the number of combinations for 30 potions.

IV. Pilot study

We recruited 30 first-year students from a local university to participate in the study, aiming to assess both the game design and the impact of EscapeDA on promoting student engagement in tutorial preparation. These students were enrolled in the module "Data Structures and Algorithms" and used the game to prepare for the tutorial of Depth First Search (DFS) and Dynamic Programming before attending. The game was released to students one week before they attend the tutorial session. After the gaming experience, students were asked to complete a questionnaire consisting of two parts: one evaluating the game design to understand their intuitive feelings and acceptance and the other assessing the learning experience to explore the actual impact and mechanisms of the game on student engagement. A total of 21 valid responses were collected. Of these participants, 61.90% were male and 38.10% were female, with an average age of 21.66 years.

Game Design Evaluation

In this part, we utilized binary opposition word pairs to capture students' intuitive reactions to various aspects of the game design [14]. To further quantify their attitudes and feelings, we employed a 7-point Likert scale for detailed assessment. The evaluation criteria included the following:

- Obstructive or Supportive: This criterion assesses how the game design influences the player's experience, focusing on whether it facilitates or hinders the achievement of game objectives.
- Complicated or Easy: This item evaluates the game's complexity and usability, reflecting its accessibility and the learning curve, which are particularly important for initial player interactions.
- Inefficient or Efficient: This criterion focuses on the efficiency of the game, particularly whether the interface and mechanics enable players to complete tasks effectively.
- Confusing or Clear: Clarity is central to game design; this item assesses whether the game effectively communicates the rules and objectives to players.
- Boring or Exciting: This item measures the entertainment value of the game, assessing its ability to sustain interest and engagement.
- Not Interesting or Interesting: This aspect examines the attractiveness of the game's content, particularly whether it sparked curiosity and encouraged exploration.
- Conventional or Inventive: This item examines the innovativeness of the game design, assessing whether the game offers a unique experience or follows industry norms.
- Usual or Leading Edge: This criterion explores whether the game incorporates cutting-edge technology or concepts, reflecting the advancement and modernity of its design.

As illustrated in Figure 4, the results show that most students responded positively to the game, particularly in terms of its supportiveness, efficiency, entertainment value, and innovation. For instance, 71% of the participants found the game innovative, indicating that the design offered an enjoyable player experience, underlying the acceptance of the puzzle design provided by LLMs. However, there is a significant divergence in evaluations of the game's complexity, clarity, and some aspects of interactivity. These findings suggest a need for further refinement to simplify gameplay, clarify objectives and rules, and enhance the game's overall accessibility and appeal, which further implies human effort is required to iteratively refine the design generated by LLMs.

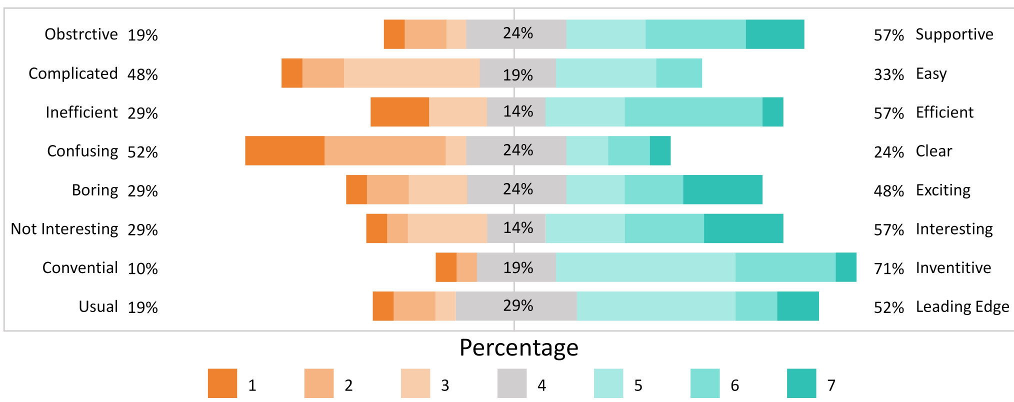


Figure 4. Game Design Evaluation

A. Engagement evaluation

The assessment of students' engagement was divided into three dimensions: behavioral, emotional, and cognitive engagement, encompassing a total of 27 questions [8]. Students rated their responses on a 5-point scale. The specifics of the questionnaire design are as follows:

- Behavioral Engagement (BE1-BE4): This dimension assesses students' behaviors, such as attentiveness and effort during gameplay, to quantify their level of participation and focus in the game.

- Emotional Engagement (EE1-EE4): By understanding students' emotional responses to the game, we assess the gamified tutorial content’s attractiveness and its effectiveness in sparking enthusiasm for learning.
- Cognitive Engagement (CE1-CE8): This dimension measures students' intrinsic thought processes during learning, such as their ability to connect knowledge, construct meaning, and engage in deep, reflective learning.

Item	Question	Mean	Std.
BE1	I listen carefully in playing game.	3.857	1.062
BE2	The first time I encounter a new topic, I listen very carefully.	4.190	0.750
BE3	I work hard when starting something new in game session.	4.238	0.944
BE4	I pay attention in game session.	4.429	0.870
EE1	I enjoy learning new things in game session.	4.523	0.680
EE2	When I work on something in game sessions, I feel interested.	4.571	0.676
EE3	When I am in game session, I feel curious about what I am learning.	4.333	0.796
EE4	The game session is fun.	3.714	1.419
CE1	When completing tasks in game, I try to relate what I’m learning to what I already know.	4.190	0.178
CE2	When I study in game, I try to connect what I am learning with my own experiences.	4.238	1.044
CE3	I try to make all the different ideas fit together and make sense when I study in game.	4.143	1.062
CE4	I make up my own examples to help me understand the important concepts I study in game.	3.714	1.189
CE5	Before I begin to study in game, I think about what I want to get done.	3.428	1.362
CE6	When I’m working on tasks, I stop once in a while and go over what I have been doing	3.952	1.203
CE7	As I study in game, I keep track of how much I understand, not just if I am getting the right answers.	3.667	1.197
CE8	If what I am working on is difficult to understand, I change the way I learn the material.	4.048	0.740

Table 1. Summary of Engagement Evaluation

As shown in Table 1, the overall average scores for behavioral, emotional, and cognitive engagement are high, indicating that participants were able to maintain focused attention, experience positive emotions, and exhibit deep psychological involvement during the game-based learning process. Moreover, participants consistently reported that the game had clear learning objectives and strong relevance to the learning content, significantly enhancing the learning experience. These findings underscore the importance of embedding well-defined learning objectives into the prompts when designing educational games with MLLMs.

V. Conclusion

In this work, we leveraged MLLMs (e.g., GPT-4V) to design and develop EscapeDA, an educational escape room game, to explore how MLLMs can reduce the resources required for serious game design and development in the higher educational context. By utilizing GPT-4V's capabilities, we demonstrated a small team with limited resources could successfully create a serious game. The pilot user study revealed that most students provided positive feedback on the game's design, reporting high levels of behavioral, emotional, and cognitive engagement during gameplay. These findings highlight that incorporating LLMs into the game design process not only aligns the game with specific learning objectives but also enhances the student learning experience by creating an engaging and interactive environment.

Looking ahead, future research will focus on refining the application of LLMs in educational game development. Specifically, we plan to develop a dedicated tool to facilitate the gamification of learning objectives and content, integrating the insights derived from this study. Additionally, we aim to conduct long-term user studies to evaluate the sustained effects of gameplay on learning outcomes and student engagement once a fully developed game is deployed.

Acknowledgments. This research is supported by EdeX Teaching and Learning Grant, Nanyang Technological University, Singapore.

References

[1] Asad Anjum, Yuting Li, Noelle Law, M. Charity, and Julian Togelius. 2023. The Ink Splotch Effect: A Case Study on ChatGPT as a Co-Creative Game Designer. *arXiv preprint arXiv:2303.02454* (2023). http://arxiv.org/abs/2303.02454.

[2] Jacob Bishop, and Matthew A Verleger. 2013 The flipped classroom: A survey of the research. In *2013 ASEE annual conference & exposition*, 23-1200.

[3] Sebastian Deterding, Dan Dixon, Rilla Khaled, and Lennart Nacke. 2011. From game design elements to gamefulness: defining "gamification:. In *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments*. 9–15.

[4] Christo Dichev and Darina Dicheva. 2017. Gamifying education: what is known, what is believed and what remains uncertain: a critical review. *International journal of educational technology in higher education* 14 (2017), 1–36.

[5] Anastasia Dimitriadou, Naza Djafarova, Ozgur Turetken, Margaret Verkuyl, and Alexander Ferworn. 2021. Challenges in serious game design and development: Educators' experiences. *Simulation & Gaming* 52, 2 (2021), 132–152.

[6] Panagiotis Fotaris and Theodoros Mastoras. 2019. Escape rooms for learning: A systematic review. In *Proceedings of the European Conference on Games Based Learning*. 235–243.

[7] Panagiotis Fotaris, Theodoros Mastoras, and Petros Lameras. 2023. Designing educational escape rooms with generative AI: A framework and ChatGPT prompt engineering guide. In *Proceedings of the European Conference on Games-based Learning*.

[8] Jennifer A Fredricks, Phyllis C Blumenfeld, and Alison H Paris. School engagement: Potential of the concept, state of the evidence. *Review of educational research* 74, no. 1 (2004): 59-109.

[9] Tommaso Fulcini and Marco Torchiano. 2023. Is ChatGPT capable of crafting gamification strategies for software engineering tasks?. In *Proceedings of the 2nd International Workshop on Gamification in Software Development, Verification, and Validation*. 22–28.

[10] Roberto Gallotta, Graham Todd, Marvin Zammit, Sam Earle, Antonios Liapis, Julian Togelius, and Georgios N Yannakakis. 2024. Large language models and games: A survey and roadmap. *arXiv preprint arXiv:2402.18659* (2024).

[11] Christopher Groening and Carmen Binnewies. 2019. “Achievement unlocked!” – The impact of digital achievements as a gamification element on motivation and performance. *Computers in Human Behavior* 97 (2019), 151–166.

[12] Jean Tate Hiebert. 2011. Gamestorming: A Playbook for Innovators, Rulebreakers, and Changemakers. *Journal of Library Innovation* 2, 2 (2011), 101–103.

[13] Jack Kelly, Michael Mateas, and Noah Wardrip-Fruin. 2023. Towards Computational Support with Language Models for TTRPG Game Masters. In *Proceedings of the 18th International Conference on the Foundations of Digital Games*. ACM, Lisbon, Portugal, 1–4.

[14] Ahmet Baki Kocaballi, Liliana Laranjo, and Enrico Coiera. Understanding and measuring user experience in conversational interfaces. *Interacting with Computers* 31, no. 2 (2019): 192-207.

[15] Vikram Kumaran, Jonathan Rowe, Bradford Mott, and James Lester. 2023. SceneCraft: Automating Interactive Narrative Scene Generation in Digital Games with Large Language Models. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, Vol. 19. 86–96.

[16] Petros Lameras, Sylvester Arnab, Ian Dunwell, Craig Stewart, Samantha Clarke, and Panagiotis Petridis. 2017. Essential features of serious games design in higher education: Linking learning attributes to game mechanics. *British journal of educational technology* 48, 4 (2017), 972–994.

[17] Pier Luca Lanzi and Daniele Loiacono. 2023. Chatgpt and other large language models as evolutionary engines for online interactive collaborative game design. In *Proceedings of the Genetic and Evolutionary Computation Conference*. 1383–1390.

[18] Joey J Lee and Jessica Hammer. 2011. Gamification in education: What, how, why bother? *Academic exchange quarterly* 15, 2 (2011), 146.

[19] Elias Melchor-Ferrer and Maria Angeles Davia-Rodriguez. 2023. Computer games and the study of terminology: An application to national accounts. *Education and Information Technologies* 28, 1 (2023), 135–153.

[20] David R Michael and Sandra L Chen. 2005. *Serious games: Games that educate, train, and inform*. Muska & Lipman/Premier-Trade.

[21] Andy Nealen, Adam Saltsman, and Eddy Boxerman. 2011. Towards minimalist game design. In *Proceedings of the 6th international conference on foundations of digital games*. 38–45.

[22] Eduardo Ortiz-Martínez, José Santos-Jaén, and Manuel Palacios-Manzano. 2022. Games in the Classroom? Analysis of Their Effects on Financial Accounting Marks in Higher Education. *International Journal of Management in Education* 20, 1 (2022), 100584.

[23] Ciprian Paduraru, Alin Stefanescu, and Augustin Jianu. 2023. Unit Test Generation Using Large Language Models for Unity Game Development. In *Proceedings of the 1st ACM International Workshop on Foundations of Applied Software Engineering for Games*. ACM, Porto de Galinhas, Brazil, 7–13.

[24] Jan L Plass, Bruce D Homer, and Charles K Kinzer. 2015. Foundations of game-based learning. *Educational psychologist* 50, 4 (2015), 258–283.

[25] Lisa-Maria Putz, Florian Hofbauer, and Horst Treiblmaier. 2020. Can Gamification Help to Improve Education? Findings from a Longitudinal Study. *Computers in Human Behavior* 110 (9 2020), 106392.

[26] Shuang Qiao, Susanna Siu-sze Yeung, Xiang Shen, and Samuel Chu. 2022. The Effects of a Gamified Morphological Awareness Intervention on Students' Cognitive, Motivational and Affective Outcomes. *British Journal of Educational Technology* 53, 4 (2022), 952–976.

[27] Shen Qiao, Susanna Siu-sze Yeung, Zamzami Zainuddin, Davy Tsz Kit Ng, and Samuel Kai Wah Chu. 2023. Examining the effects of mixed and non-digital gamification on students' learning performance, cognitive engagement and course satisfaction. *British Journal of Educational Technology* 54, 1 (2023), 394–413.

[28] Mohammad Reza Taesiri, Finlay Macklon, Yihe Wang, Hengshuo Shen, and Cor-Paul Bezemer. 2022. Large Language Models Are Pretty Good Zero-Shot Video Game Bug Detectors. *arXiv preprint arXiv:2210.02506* (2022).

[29] Sigmund Tobias, J Dexter Fletcher, and Alexander P Wind. 2014. Game-based learning. *Handbook of research on educational communications and technology* (2014), 485–503.

[30] Graham Todd, Sam Earle, Muhammad Umair Nasir, Michael Cerny Green, and Julian Togelius. 2023. Level Generation Through Large Language Models. In *Proceedings of the 18th International Conference on the Foundations of Digital Games*. ACM, 1–8.

[31] Judith van Stegeren and Jakub Myśliwiec. 2021. Fine-Tuning GPT-2 on Annotated RPG Quests for NPC Dialogue Generation. In *Proceedings of the 16th International Conference on the Foundations of Digital Games (FDG '21)*. ACM, New York, NY, USA.

[32] Susanna Värtinen, Perttu Hämäläinen, and Christian Guckelsberger. 2022. Generating Role-Playing Game Quests with GPT Language Models. *IEEE Transactions on Games* (2022), 1–12.

[33] Soumia Yamoul, Lynda Ouchaouka, and Mohamed Moussetad. 2024. Towards a Wider Adoption of Serious Games in Higher Education: Deepening the Analysis of Benefits and

Challenges. *International Journal of Information Science and Technology* 8, 4 (2024), 16–25.

Appendix

A. *Tutorial Questions to be Gamified*

Question 1. Give a pseudocode of finding a simple path connecting two given vertices in an undirected graph by Depth-First-Search.

Question 2. Give a pseudocode of finding the Fibonacci number for a given n by using Dynamic Programming strategy.

B. *An example of prompt*

You are a gamification designer, provide a gamification idea for the given tutorial question for the first-year students enrolled in the module of Data Structure and Algorithm. The idea needs to stick with the learning objectives, fits the given theme and is practical to implement. Use of what available in the given environment as much as possible.

Question: Give a pseudocode of finding a simple path connecting two given vertices in an undirected graph by Depth-First-Search.

Learning Objectives: Deepen students' understanding of the Depth-First-Search Algorithm by practicing the algorithm.

Theme: