# Student Engagement and Perceived Learning in an Educational Platform Game with Automatic Question Generation: A Brazilian High School Study

**Matheus Prado Prandini Faria**[1], **Rita Maria Silva Julia**[1], **Marcelo Zanchetta do Nascimento**[1], **Arthur Henrique de Souza**[1], **Bruno Castro Lima**[1]

[1]Computer Science Department – Federal University of Uberlândia
Uberlândia, Brazil

{matheus_prandini,rita,marcelo.nascimento}@ufu.br

{arthurhsouza,brunocastro}@ufu.br

***Abstract. Introduction:*** *Digital games show promise for engaging students in mathematics, but few studies examine educational games with automatic question generation in real high school contexts.* ***Objective:*** *This paper investigates student and teacher perceptions of Gael, a 2D educational platform game with multiple-choice question generation, for teaching financial mathematics in 11th grade (second year of Brazilian high school).* ***Methodology:*** *An exploratory case study involved 61 students from a Brazilian public high school, using a post-session questionnaire with ten five-point Likert items on gameplay, engagement, content adequacy, and perceived learning, plus three open-ended questions. Teacher feedback and game logs on in-game performance served as complementary data.* ***Results:*** *Students reported high fun, attention, usability, and intention to reuse the game, while perceived learning was moderately positive and varied. Student and teacher comments identified gameplay, aesthetics, and integration of financial mathematics as strengths, and mathematical difficulty, limited prior knowledge, and control issues as main challenges.*

***Keywords*** *Educational games, Game-based learning, Automatic question generation, Student engagement, Financial mathematics*

## 1. Introduction

Digital educational games and game-based learning (GBL) have been widely studied as ways to increase students' motivation and engagement in mathematics. Empirical studies and recent reviews report positive effects of GBL on interest, attitudes, enjoyment, and often on cognitive outcomes and achievement in K–12 mathematics [Hui e Mahmud 2023, Li et al. 2024b]. In Brazilian secondary education this domain is addressed as *financial education* (*educação financeira*), a cross-cutting curricular theme, rather than as *financial mathematics* in the higher-education sense; here we target the mathematical content that underpins it (percentage and simple and compound interest). It has grown in secondary curricula for its role in everyday decision making and citizenship, and several games target arithmetic, proportional reasoning, and financial concepts in basic education, frequently reporting higher engagement

and more contextualized problem solving [Guimarães et al. 2024, Andrade et al. 2025, Vieira et al. 2025b, Vieira et al. 2025a]. However, many of these games use fixed question sets and scripted scenarios, limiting adaptivity, and teachers' ability to align content with local curricula and evolving learning goals.

Advances in automatic question generation (AQG) now enable dynamic creation of practice items from textual or structured learning materials, especially in mathematics and other STEM domains [Rockembach e Thom 2026, Li et al. 2024a]. Recent work explores combining AQG with intelligent tutoring systems and adaptive practice platforms, including generative AI approaches that produce mathematics questions with controlled difficulty and cognitive depth [Song et al. 2025, Yu et al. 2025]. Yet most AQG research focuses on standalone formative assessment banks or tutoring systems, with little integration into full GBL environments where questions must fit narrative gameplay and moment-to-moment player actions. Existing surveys find scarce empirical evidence on AQG within complete GBL environments, particularly in authentic classroom settings [Rockembach e Thom 2026, Li et al. 2024a], a gap that is especially pronounced in Brazilian high school mathematics education [Sousa et al. 2025].

Beyond AQG, evaluating educational games remains methodologically challenging: frameworks emphasize combining multiple sources of evidence—self-reports, gameplay logs, and qualitative feedback—to capture both engagement and learning outcomes [Hui e Mahmud 2023, Gresse von Wangenheim et al. 2025], and learning gains depend on how well tasks align with curricular objectives and how teachers integrate games into practice [de Melo et al. 2021, Russo et al. 2025, Pesout et al. 2025].

Motivated by these gaps, this paper presents Gael, a 2D educational platform game that can be parametrized with different curricular contents and is coupled with a multiple-choice question (MCQ) AQG module. The study reports an exploratory case study of Gael's deployment in financial mathematics lessons with 11th-grade students in a Brazilian public high school, combining student questionnaires, teacher interviews, and gameplay logs. The contributions of this work are threefold: (i) an educational platform-game architecture integrated with a question-generation backend, instantiated for financial mathematics; (ii) a multi-perspective evaluation protocol triangulating student self-reports, teacher feedback, and behavioral logs in an authentic high-school context; and (iii) empirical evidence on engagement, perceived pedagogical value, and practical deployment challenges in Brazilian high-school mathematics education. Specifically, the study addresses the following research questions: **(RQ1)** How do students perceive their engagement and gameplay experience when using Gael in financial mathematics lessons? **(RQ2)** How do students and teachers perceive the adequacy and contribution of Gael and its automatically generated questions to reviewing and learning financial mathematics? **(RQ3)** What do gameplay logs reveal about students' in-game performance and interaction patterns?

## 2. Related Work

Systematic reviews consistently report positive effects on motivation, enjoyment, and often achievement [Hui e Mahmud 2023, Li et al. 2024b]. In financial mathematics, for instance, games like MathCity embed percentage and proportional reasoning in urban planning scenarios, increasing engagement and contextualized problem-

solving [Guimarães et al. 2024]. A review of financial literacy games found that most interventions improved self-reported financial knowledge and attitudes, though few assessed whether these gains transferred to real-world financial decisions [Andrade et al. 2025].

Despite progress, many mathematics games still use fixed question sets authored during development, limiting reuse across curricula and restricting teachers' ability to align content with local learning objectives [de Melo et al. 2021]. Empirical studies of game-based learning in authentic Brazilian high school classrooms are also scarce. A systematic mapping of Brazilian educational games for exact sciences found that 76% were not publicly available, hindering school adoption [Sousa et al. 2025]. Few studies systematically included teacher perspectives or reported real classroom deployments, especially in mathematics [Russo et al. 2025, Pesout et al. 2025]. This gap motivates the present study's focus on classroom-based evaluation with teacher feedback.

AQG has emerged as a complementary approach to scale content in intelligent tutoring systems and adaptive learning platforms. Recent surveys group AQG into template-based, rule-based, and neural (LLM) methods [Li et al. 2024a, Rockembach e Thom 2026], and show that LLM-generated math questions can match human-authored items in correctness and difficulty when paired with validation pipelines: GPT-4o has produced math questions with high appropriateness and correctness [Yu et al. 2025], and GPT-4 has improved semantic coherence in contextualized problems [Li et al. 2024a]. A review of automatic item generation underscores its potential to reduce instructor workload while maintaining quality [Song et al. 2025]. However, most AQG work targets standalone formative-assessment pools, with limited integration into full GBL environments where questions must be embedded within narrative gameplay [Rockembach e Thom 2026, Li et al. 2024a].

Evaluating educational games is methodologically challenging, as engagement and learning outcomes interact. Frameworks recommend multi-method approaches that triangulate self-reports, behavioral logs, and qualitative feedback [Hui e Mahmud 2023, Gresse von Wangenheim et al. 2025]; models such as MEEGA+ offer standardized questionnaires for player experience, usability, and perceived learning [Giani Petri e Borgatto 2019]. Studies also stress teacher perspectives, since teachers mediate classroom integration and long-term adoption [Russo et al. 2025, Pesout et al. 2025]: students may rate games highly for enjoyment, while teachers report curriculum misalignment and limited progress tracking as barriers [Russo et al. 2025]. Recent work further highlights metacognitive scaffolding—prompts linking gameplay to learning objectives—as key to narrowing the engagement–learning gap [Cheng et al. 2017, Vidaković et al. 2023].

## 3. Gael: An Educational Platform Game with Automatic Question Generation

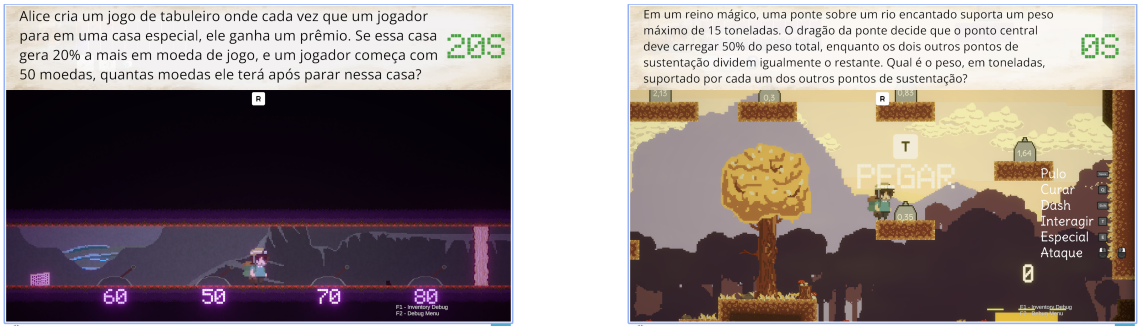
This section describes the design and implementation of Gael (Contribution i), an educational 2D platform game integrated with an automatic question generation backend.

### 3.1. Game Design and Mechanics

Gael is a 2D action-platformer developed in Unity and deployed as a WebGL application accessible via web browsers. The player controls the protagonist with core mechanics including jumping, dashing, and combat using melee (sword) and ranged (bow) attacks.

**Level Archetypes and Progression.** The game implements an infinite progression system with six randomly sampled level archetypes, ensuring variety and unpredictability. Levels are selected uniformly at random from the pool, with no adaptive sequencing or difficulty progression in the current version. The six archetypes are: (1) **lever-based MCQ levels** where players activate the lever corresponding to the correct answer among four alternatives to unlock the exit door; (2) **weight-based calculation levels** requiring numerical answers by collecting and combining weighted objects whose sum equals the solution; (3) **combat-only levels** with sequential enemy waves that must be defeated; (4) **color memory sequences** where players replicate a progressively longer pattern; (5) **mathematical grid puzzles** where players arrange numbers to maximize one quadrant's sum and minimize another's; and (6) **logic puzzles**, including construction tasks, lever-combination challenges, and "river-crossing" puzzles inspired by the classic "Wolf, Goat, and Cabbage" problem.

**Feedback and Rewards.** Each level awards 0–3 stars based on completion time, providing immediate performance feedback. The heads-up display (HUD) shows the player's health bar, collected coins (currently cosmetic) and elapsed time. When players enter question-based levels (archetypes 1 and 2), questions appear as full-screen overlays while gameplay continues; players must navigate the environment and answer the question simultaneously, creating a dual cognitive-motor challenge (see Figure 1).




(a) Lever-based MCQ level

(b) Weight-based calculation level

**Figure 1. Examples of Gael's level archetypes, showing (a) a MCQ presented via levers and (b) a numerical puzzle using weighted objects**

### 3.2. Automatic Question Generation Module

The question generation backend is implemented as a RESTful API that leverages OpenAI's GPT-4o model to produce contextualized MCQ tailored to specific curricular content. GPT-4o was chosen for its ability to generate semantically coherent distractors and to reframe mathematical problems into narrative-driven contexts that align with the game's fantasy setting.

**Generation Pipeline.** The pipeline has four stages: (1) **Document ingestion:** educators upload PDFs or text documents which are preprocessed; the system extracts

and preprocesses the content. (2) **Question generation:** the preprocessed text is sent to GPT-4o through a single structured prompt that instructs the model to derive four-option MCQs (one correct answer) from the ingested content, reframe each problem within the game's fantasy narrative (for example, percentage calculations as a wizard's spell cost or a merchant's price), and return a JSON object with the question text, the four alternatives, the correct-answer index, and topic and difficulty tags. (3) **Question validation:** questions undergo (a) an LLM-based check for mathematical correctness, clarity, and alignment with the source, and (b) a code-based check for structural integrity (exactly four options, one correct answer, no empty fields). In the financial mathematics setting, about 20% of questions were rejected for ambiguity or incorrect calculations. Validation was fully automated, with no manual review by a subject-matter expert (revisited in Section 6). (4) **Storage and retrieval:** validated questions are stored in Firestore with metadata (topic, difficulty, content source) and can be reused across students in the same grade and school.

**Distractor Design.** Incorrect alternatives are generated by the LLM using two strategies: producing plausible numerical values within contextually appropriate ranges and anticipating common student misconceptions (for example, confusing simple and compound interest formulas, applying percentage operations incorrectly).

### 3.3. Integration and Instantiation for Financial Mathematics

**Game-Backend Communication.** The game communicates with the backend via RESTful HTTP requests. When a player enters a question-based level, the game sends a GET request specifying the content domain ("financial mathematics"); the API returns a JSON object containing the question text, four alternatives, the correct answer index, and metadata (question ID, topic tag). Additionally, each time a player submits an answer, a POST request logs the response (correctness, time taken) to support learning analytics.

**Question Generation Timing.** An initial pool of 20 questions is pre-generated upon document ingestion and stored in the database. If the pool is exhausted during gameplay or if an educator requests additional items, new questions are generated on demand via a manual trigger in the admin interface. In this study, all questions were generated and validated offline before the sessions; gameplay retrieved only pre-validated items, with no LLM calls in real time.

**Content Instantiation.** In this case study, Gael was instantiated with content aligned to BNCC competency **EM13MAT203** (applying mathematical concepts to everyday financial decisions, such as household budget control and interest calculators). The question pool covered three financial mathematics topics: percentage, simple interest, and compound interest, usually framed in everyday financial scenarios (such as investments, discounts and purchases) and embedded in gamified narratives. In the 20-item pool, 60% of questions covered percentage calculations, 20% simple interest, and 20% compound interest. Questions were tagged with school and grade metadata and shared within the same context to ensure consistent content. Table 1 shows two representative examples.

## 4. Study Design

This section describes the exploratory case study conducted to evaluate Gael in authentic classroom settings (Contribution ii). The study received ethical approval and all

**Table 1. Example questions generated for financial mathematics**

| |
|---|
| **Q1 (Simple interest):** *The inventor of the future wants to triple his invention by investing R$ 1,000 at a simple interest rate of 10% per year. How long, in years, does he need to wait to triple his investment?* **Correct:** 20 years |
| **Q2 (Compound interest):** *Princess Juliana decided to save R$ 600 in her Enchanted Cave account, which yields 2% per month in compound interest. After 1 year, what will be the total amount Juliana has in her account?* **Correct:** R$ 760.24 |

participants provided informed consent[1].

### 4.1. Context and Participants

The study took place at a public state high school [location omitted] with three 11th-grade classes (N=97). Students who missed any of the three gameplay sessions were excluded from analysis, yielding a final sample of N=61 (mean age=16.15, SD=0.36). The students had no prior exposure to financial mathematics or Gael. The school administration requested the deployment as preparation for a financial literacy competition, not as part of a planned curricular unit. Two mathematics teachers and three researchers attended all sessions to provide technical support, mediate gameplay, and observe interactions, but did not pre-teach mathematical content.

### 4.2. Procedure

The intervention comprised three 50-minute sessions over three weeks (150 minutes per student) in the school computer lab, where each student used a desktop via a web browser. Teachers and research staff briefly introduced game controls but did not pre-teach financial mathematics. During gameplay, they resolved technical issues, clarified mechanics, and occasionally provided hints when students were clearly stuck on math questions to balance challenge and frustration. Students progressed at their own pace through randomly selected levels.

### 4.3. Data Collection Instruments

Data were collected from three sources to allow triangulation.[2]

**Student Questionnaire:** Ten Likert-scale items (1=Strongly Disagree to 5=Strongly Agree) measuring: (i) Fun/Enjoyment, (ii) Usability/Clarity, (iii) Attention/Motivation, and (iv) Perceived Learning/Reuse Intention. Three open-ended questions elicited feedback on likes, dislikes and perceived learning contribution. Administered after the final session.

**Teacher Questionnaire:** Seven Likert items assessing engagement, alignment, feasibility, motivation, learning support and future use potential, plus three open-ended questions on strengths, challenges and suggestions.

[1]Process number 78520523.0.0000.5152 approved by the Comitê de Ética em Pesquisas com Seres Humanos at the Federal University of Uberlândia (CEP-UFU).

[2]In the interest of open science, the two questionnaires and the anonymized response data are available at `https://github.com/matheus-prandini/adapt2learn-research`.

**Gameplay Logs:** Automatic logging of in-game events, such as time per level, answer correctness, attempts, stored in a database.

### 4.4. Data Analysis

**Quantitative Analysis (RQ1, RQ2):** Likert responses were analyzed using descriptive statistics (mean, SD, median). Given the ordinal nature of the data and non-normal distributions, medians and interquartile ranges were prioritized.

**Qualitative Analysis (RQ2):** Open-ended responses were thematically coded to identify recurring themes related to strengths (e.g., gameplay, aesthetics), challenges (e.g., question difficulty, controls) and perceived learning contributions.

**Log Analysis (RQ3):** Per-student metrics computed included: levels completed, average time per level, proportion of correct first attempts, attempts per question and average star rating. Performance was disaggregated by question topic (percentages, simple interest, compound interest) to identify content-specific patterns. All analyses were conducted using Python (pandas, matplotlib).

## 5. Results and Discussion

This section presents findings (Contribution iii) organized by RQ, triangulating evidence from student questionnaires, teacher feedback, and gameplay logs.

### 5.1. Student Engagement and Gameplay Experience (RQ1)

Table 2 summarizes responses to the ten Likert-scale items (N=61). Students reported high levels of fun (M=4.23, SD=0.78), attention/concentration (M=4.20, SD=0.85) and intention to reuse the game (M=4.05, SD=1.13). Usability ratings were moderately positive: clarity of objectives (M=3.95, SD=0.86) and ease of controls (M=4.03, SD=0.86) were both above the midpoint. Motivation to continue playing in class was also high (M=3.97, SD=1.17). On the other hand, perceived Learning (M=3.75, SD=1.14) showed the greatest variability, suggesting heterogeneous views of educational value.

Qualitative feedback (N=61 open-ended responses) revealed that students most appreciated **gameplay mechanics** (49% of responses), particularly combat and platforming challenges (for example, "killing the monsters", "jumping between platforms"). **Aesthetics and design** were mentioned by 12% of students (for example, "the cute character", "pixel art graphics"). **Learning-related aspects** were highlighted by 18%, including the gamified presentation of questions and opportunities for reasoning (for example, "helped me think more about calculations").

The high fun (M=4.23) and concentration (M=4.20) alongside moderate perceived learning (M=3.75) suggest a gap between engagement and metacognitive awareness: students were immersed in gameplay but uncertain of learning gains, consistent with research showing immersion does not always predict perceived outcomes [Cheng et al. 2017], which may also reflect unfamiliarity with the content.

### 5.2. Perceived Adequacy and Learning Contribution (RQ2)

#### 5.2.1. Student Perceptions

When asked how Gael contributed to learning financial mathematics, 56% of students indicated the game helped in general terms, often mentioning improved understanding of

**Table 2. Student questionnaire responses (N=61)**

| Item / Dimension | Mean | SD | Median |
|---|---|---|---|
| *Fun & Enjoyment* | | | |
| The game was fun to play | 4.23 | 0.78 | 4.0 |
| *Usability & Clarity* | | | |
| I understood what to do | 3.95 | 0.86 | 4.0 |
| Commands/controls were easy | 4.03 | 0.86 | 4.0 |
| *Content Quality* | | | |
| Reasoning challenges were interesting | 4.00 | 0.84 | 4.0 |
| Math questions made sense | 3.92 | 0.92 | 4.0 |
| Difficulty level was adequate | 3.92 | 0.82 | 4.0 |
| *Attention & Engagement* | | | |
| I stayed focused/concentrated | 4.20 | 0.85 | 4.0 |
| *Perceived Learning* | | | |
| The game helped me learn content | 3.75 | 1.14 | 4.0 |
| *Motivation & Reuse* | | | |
| I felt motivated to keep playing | 3.97 | 1.17 | 4.0 |
| I would use Gael again | 4.05 | 1.13 | 4.0 |

percentage calculations (8% specifically). Representative quotes include:

> *"Helped me understand percentage better"* (Student 9)
> *"Turned math from boring worksheets into a way to level up."* (Student 18)
> *"Helped me review content I had forgotten"* (Student 43)

However, 5% of students said the game **did not help much**, citing lack of prior knowledge or difficulty with the questions (e.g., "I never studied this content, so it didn't help"). Three students said the game mainly exposed gaps in their mathematical knowledge rather than addressing them (e.g., "made me see I'm not good at this yet").

The main challenge was **question difficulty** (33%), with students citing complex calculations and limited time to achieve the highest star rating (e.g., "some questions were really hard and response time was too short"). **Control issues** involving hand skill were reported by 23% (for example, "jumping and dashing simultaneously was hard"). The **lack of persistent checkpoints** was noted by 5% (e.g., "having to restart from the beginning every session was frustrating"), indicating a clear design improvement.

### 5.2.2. Teacher Perspectives

Both teachers rated Gael positively on engagement (M=4.5), alignment with grade level (M=4.0), feasibility (M=4.5), and future use potential (M=4.5). However, perceived student motivation compared to traditional activities received divergent ratings (Teacher 1: 2, Teacher 2: 4, M=3.0), and perceived learning support was moderate (M=3.5).

Qualitative feedback highlighted **strengths**: "change of classroom dynamics," "ludic approach", "well-designed logic problems", and "helped students understand percentage". **Challenges** focused on content misalignment: because financial mathematics was not in the planned curriculum, students lacked prior exposure, which

"hindered learning and engagement", according to Teacher 1. Teacher 2 observed that although some content was unfamiliar, "students sought out information and managed to understand."

A key suggestion was to improve the **onboarding experience**: Teacher 1 recommended an introductory session explaining the game's objectives and how mathematical questions integrate with gameplay, to clarify the activity's purpose. Teacher 2 suggested expanding Gael to other mathematical topics beyond financial math.

The curriculum misalignment represents a critical limitation: when students lack foundational knowledge, games risk becoming frustrating rather than facilitative [Cai et al. 2026], and Teacher 1's onboarding suggestion aligns with research showing that explicit metacognitive scaffolding helps students connect gameplay to learning objectives [Vidaković et al. 2023].

### 5.3. In-Game Performance and Interaction Patterns (RQ3)

Gameplay logs from question-based levels were analyzed for all 61 students to examine behavioral indicators of engagement and mathematical performance. Because the game's infinite progression system randomly samples different level archetypes, the present paper focused on levels with mathematical questions (lever-based and weight-based), which provide direct evidence of content engagement.

**Question Engagement and Correctness.** On average, students answered 6.05 mathematical questions (SD=3.19) across three sessions. This low number reflects the game's design, where lever-and-weight questions are randomly mixed with combat and logic puzzles, producing variable mathematical exposure. First-attempt accuracy was 77.7% (SD=17.87%), indicating generally high success but substantial individual variation. Because a correct answer is required to unlock the exit and advance, nearly all students (99%) continued attempting items after errors until they responded correctly.

**Performance Across Sessions.** First-attempt correctness declined from Session 1 (M=85.22%) to Session 3 (M=76.16%), but a chi-square test showed no significant difference ($\chi^2 = 3.09$, $p = 0.079$), suggesting this change reflects random sampling from the 20-item pool rather than systematic performance shifts. In contrast, average time per question fell from 10.18 minutes to 6.23 minutes, indicating more efficient navigation and problem-solving. Overall, students used the game more quickly, but limited question exposure (M=6.05 questions per student) and random item distribution did not produce measurable gains in mathematical accuracy.

**Performance by Content Topic.** Questions involving **percentage calculations** had the highest correctness rate (81.09%, $N$ = 312), aligning with teacher observations that this was the most accessible content area and with its larger share of the item pool. **Compound interest** questions showed a moderately high success rate (72.00%, $N$ = 25), despite both students and teachers reporting this as the most challenging and unfamiliar topic. **Simple interest** questions had the lowest correctness rate (63.27%, $N$ = 49), possibly reflecting conceptual confusion between simple and compound interest formulas and fewer opportunities to stabilize solution strategies for each type of problem. $N$ denotes the total number of responses.

**Time on Task.** Across all sessions, question-based levels averaged 8.14 minutes

per question (SD=8.11), with substantial variation due to question complexity and students' mathematical fluency.

Faster performance without better accuracy ($p = 0.079$) suggests students improved procedural fluency in game mechanics but not mathematical understanding. Low item exposure (M=6.05) was below deliberate-practice levels, which require many focused, feedback-rich attempts on specific sub-skills, and was likely exacerbated by students' lack of prior formal instruction in financial mathematics. Overall, despite limited mathematical gains under current conditions, Gael was well accepted and effectively supported engagement in this exploratory classroom deployment.

## 6. Conclusion and Future Work

This paper evaluated Gael, a 2D platformer with an LLM-based AQG module for financial mathematics, in a Brazilian high school classroom, examining engagement (RQ1), perceived learning (RQ2) and performance (RQ3). Students reported high engagement, fun, concentration and reuse intention, but only moderate perceived learning, indicating a gap between immersion and metacognitive awareness. Teachers judged the approach feasible but noted that introducing financial mathematics before formal instruction reduced effectiveness due to missing prerequisites. Gameplay logs showed limited mathematical exposure and faster navigation without significant accuracy gains, suggesting gains in procedural fluency rather than conceptual understanding.

This work provides empirical evidence that GBL has potential to sustain engagement in mathematics while highlighting key design factors for effective learning: curriculum alignment, sufficient question exposure and metacognitive scaffolding. As an exploratory study, it captures perceived learning and in-game performance rather than objective learning gains; without a pre/post-test and a control group, it does not support causal claims about knowledge acquisition or superiority over traditional exercises or other games. Further limitations include a small question pool (N = 20), low exposure due to random level sampling, fully automated (non-expert) question validation, teacher feedback from only two teachers, and limited responses for some topics (for example, N = 25 for compound interest). Future work should enlarge the question pool with adaptive difficulty, align deployment with the school curriculum, add metacognitive scaffolding, incorporate expert validation of generated items, and run controlled, longer interventions with pre/post assessment and comparison groups to evaluate effective learning. Overall, the study shows that closing the engagement–learning gap in game-based mathematics hinges on content alignment, exposure thresholds and metacognitive support.

## Acknowledgements

We thank our partner school, Escola Estadual Messias Pedreiro (EEMP), Uberlândia, Brazil, for the institutional support to validate our game-based educational system; the National Council for Scientific and Technological Development (CNPq), Grant #302833/2025-0; the ADASYS Game Club (Computer Science Faculty, Federal University of Uberlândia) for assistance in deploying it at the school; and Professor Stéphane Julia (same faculty) for the administrative process related to the outreach activities. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES).

## References

Andrade, S., Komati, K., Silva, E., e Bortolon, P. (2025). O uso de jogos na educação financeira: Uma revisão sistemática. In *Anais da X Escola Regional de Informática do Espírito Santo*, pages 41–50, Porto Alegre, RS, Brasil. SBC.

Cai, Z., Liu, C., Yang, Y., e Li, B. (2026). The impact of learning supports in digital game-based learning on learners with different levels of prior knowledge. *The Internet and Higher Education*, 68:101044.

Cheng, M.-T., Lin, Y.-W., She, H.-C., e Kuo, P.-C. (2017). Is immersion of any value? whether, and to what extent, game immersion experience during serious gaming affects science learning. *British Journal of Educational Technology*, 48(2):246–263.

de Melo, V., da Silva, A., Pimentel, F., Silva, R., e da Silva Júnior, L. C. (2021). Percepção e desafios na avaliação de jogos digitais educacionais. In *Anais Estendidos do XX Simpósio Brasileiro de Jogos e Entretenimento Digital*, pages 713–716, Porto Alegre, RS, Brasil. SBC.

Giani Petri, C. v. W. e Borgatto, A. (2019). Meega+: Um modelo para a avaliação de jogos educacionais para o ensino de computação. *Revista Brasileira de Informática na Educação*, 27(03):52–81.

Gresse von Wangenheim, C., Petri, G., e Ferreti Borgatto, A. (2025). Using meega/meega+ for the evaluation of educational games: A retrospective. *Revista Brasileira de Informática na Educação*, 33:943–971.

Guimarães, S., Bastos, F., Álvaro Oda, Inocêncio, A. C., Ribeiro, M. W., Névoa, M. A., e Sousa, P. M. (2024). Mathcity finances: Jogo digital educativo para praticar operações matemáticas no cotidiano. In *Anais do XXXV Simpósio Brasileiro de Informática na Educação*, pages 578–590, Porto Alegre, RS, Brasil. SBC.

Hui, H. B. e Mahmud, M. S. (2023). Influence of game-based learning in mathematics education on the students' cognitive and affective domain: A systematic review. *Frontiers in Psychology*, Volume 14 - 2023.

Li, R., Wang, Y., Zheng, C., Jiang, Y.-H., e Jiang, B. (2024a). Generating contextualized mathematics multiple-choice questions utilizing large language models. In Olney, A. M., Chounta, I.-A., Liu, Z., Santos, O. C., e Bittencourt, I. I., editors, *Artificial Intelligence in Education. Posters and Late Breaking Results, Workshops and Tutorials, Industry and Innovation Tracks, Practitioners, Doctoral Consortium and Blue Sky*, pages 494–501, Cham. Springer Nature Switzerland.

Li, Y., Chen, D., e Deng, X. (2024b). The impact of digital educational games on student's motivation for learning: The mediating effect of learning engagement and the moderating effect of the digital environment. *PLOS ONE*, 19(1):1–21.

Pesout, O., Vondráčková, V., Bačová, K., e Urban, K. (2025). Digital educational games in the eyes of educators: Teachers do not view games like us. In *ICERI2025 Proceedings*, 18th annual International Conference of Education, Research and Innovation, pages 2143–2149. IATED.

Rockembach, G. R. e Thom, L. H. (2026). The last decade of automatic question generation: A review of techniques, limitations, and applications in business process management education. *Revista Brasileira de Informática na Educação*, 34:219–251.

Russo, J. A., Roche, A., Russo, T., e Kalogeropoulos, P. (2025). Examining primary school educators' preferences for using digital versus non-digital games to support mathematics instruction. *International Journal of Mathematical Education in Science and Technology*, 56(9):1778–1803.

Song, Y., Du, J., e Zheng, Q. (2025). Automatic item generation for educational assessments: a systematic literature review. *Interactive Learning Environments*, 33(9):5386–5405.

Sousa, M., Seabra, R., e Lima, P. (2025). Educational games for teaching the essentials of exact sciences in brazil: A ten-year systematic mapping study. In *Anais do XXIV Simpósio Brasileiro de Jogos e Entretenimento Digital*, pages 1107–1124, Porto Alegre, RS, Brasil. SBC.

Vidaković, M., Lara, M., Duchi, L., Whitcomb, A., e Paas, F. (2023). Two-part onboarding for game-based learning environments. *Frontiers in Education*, Volume 8 - 2023.

Vieira, R. M., da Silva, A. P., e Damasceno, E. F. (2025a). A metaverse-based approach for financial literacy in brazilian vocational school. *Metaverse*, 6(1):3222.

Vieira, R. M., Machado, T. A., da Silva, A. P., e Damasceno, E. F. (2025b). Enhancing financial literacy by roblox: An innovative metaverse-based educational framework. In *EDULEARN25 Proceedings*, 17th International Conference on Education and New Learning Technologies, pages 53–58. IATED.

Yu, Y., Krantz, A., e Lobczowski, N. G. (2025). From recall to reasoning: Automated question generation for deeper math learning through large language models. In Cristea, A. I., Walker, E., Lu, Y., Santos, O. C., e Isotani, S., editors, *Artificial Intelligence in Education*, pages 414–422, Cham. Springer Nature Switzerland.