International Journal of Serious Games

ISSN: 2384-8766
https://journal.seriousgamessociety.org/

Article

# Gamification in Physics Education: Play Your Way to Better Learning

Katharina Richter [1], Michael Kickmeier-Rust[1]

*1 St.Gallen University of Teacher Education, Institute of Educational Psychology;, St. Gallen, Switzerland, {katharina.richter, michael.kickmeier} @phsg.ch*

**Keywords:**

Gamification
Physics Education
Knowledge Pace Theory (KST)
Learning Performance
Educational Games
Motivation
Cognitive Diagnostic Models

Received: August 2024
Accepted: February 2025
Published: February 2025
DOI: 10.17083/ijsg.v12i1.858

## Abstract

This study explores the impact of gamification on student engagement, motivation, and learning outcomes in physics education, emphasizing structural performance modeling through Knowledge Space Theory (KST). A basketball-themed educational game, incorporating gamification elements such as scoring systems, time limits, and competition, was used to teach physics concepts like initial velocity, motion, and trajectory. Using a crossover design, 10th-grade students alternated between gamified and non-gamified versions of the game. Engagement was assessed through behavioral indicators (e.g.time on tasks, number of attempts), motivation was measured using a custom questionnaire evaluating intrinsic/extrinsic factors, interest in physics, and satisfaction. Learning outcomes were evaluated through quiz scores on physics concepts before and after gameplay. Results indicate that gamification significantly increased engagement metrics and positively influenced motivation, particularly when experienced before the non-gamified condition. While quiz performance did not differ significantly between conditions, a positive correlation between game scores and quiz performance emerged during the second session. Structural performance modeling revealed the dynamic nature of learning trajectories in gamified environments. These findings highlights the potential of gamification to transform physics education by enhancing student motivation and engagement, thereby promoting a deeper understanding and retention of complex scientific concepts.

## 1. Introduction

### 1.1 Experimental Learning in Physics

The foundation of learning and research in physics is deeply intertwined with experimentation. Experimental activities are essential in the process of acquiring knowledge and understanding natural phenomena. Through experiments, scientists can observe, measure, and analyze physical phenomena to formulate hypotheses, test theories, and advance scientific knowledge.

Experimental investigations have played a pivotal role in numerous breakthroughs in physics, from Galileo Galilei's foundational experiments on motion and free fall to the discovery of new particles like the Higgs boson and the confirmation of theoretical predictions such as solar neutrino oscillations [1], [2]. Galilei's work, particularly his experiments with inclined planes and pendulums, demonstrated that all objects, regardless of mass, fall at the same rate in a vacuum. These experiments debunked Aristotle's long-held belief that heavier objects fall faster than lighter ones and laid the groundwork for Newtonian mechanics. By meticulously measuring the acceleration of falling objects and the motion of pendulums, Galilei established key principles of kinematics that continue to underpin modern physics. These experiments not only validate theoretical frameworks but also push the boundaries of scientific knowledge, leading to paradigm shifts and transformative advancements in the field of physics [3].

### 1.2 Related Works: The importance of experimental learning

Experimental setups in physics are vital for enabling active participation and inquiry-based learning, particularly in educational settings. By engaging in hands-on experiments, students can develop critical thinking skills, enhance problem-solving abilities, and deepen their understanding of scientific concepts [4] [5]. Furthermore, experiments serve as effective tools for illustrating abstract theories, fostering curiosity, and promoting a practical understanding of complex physical phenomena [6].

### 1.3 The Emergence of Gamification in Education

In recent years, gamification and serious games have gained recognition as effective tools for increasing student engagement, motivation, and overall learning experiences. Gamification is defined as the use of game design elements and mechanics in non-game contexts to increase motivation and engagement [7].
Engagement in the learning process is characterized by active and focused interest, involving cognitive, emotional, and behavioral dimensions. It is crucial for effective learning and is often linked to improved academic performance and retention [8], [9 ].

Incorporating of game elements can make learning more interactive and enjoyable, thereby fostering a deeper understanding and retention of knowledge. This approach is particularly promising in the field of physics education, where it can transform abstract concepts and complex phenomena into interactive and enjoyable learning activities. By using serious games or gamified learning applications, students can actively experiment, explore physical laws, and deepen their understanding through practical applications [10] [11]. Additionally, gamification can increase students' interest by fostering positive behavioral changes and enhancing learning motivation [11] [12].

Gamification offers a promising solution to these challenges by making learning more interactive and entertaining [13]. The integration of gamification or serious games into the field of physics learning and research can be an effective method to enhance engagement, motivation, and the overall learning experience. While gamification shows potential to improve engagement and learning outcomes, its effectiveness depends on various factors, and more research is needed to understand its long-term impact. Dichev and Dicheva [14] provide a critical review of gamification in education, emphasizing the need to understand its impact on student performance and engagement. This critical perspective highlights that, despite promising approaches, the efficiency and actual benefits of gamification still require further investigation. By incorporating gaming elements into educational practices, educators can create more interactive and engaging learning experiences that motivate students and promote long-term academic success. This is supported by numerous studies that have shown that gamification not only improves immediate learning outcomes and student performance, but also increases knowledge retention in the long term [15]. Gamification has been associated with improved learning performance and intrinsic motivation in education, suggesting that it has the potential to enhance students' understanding and mastery of specific academic content [16].

### 1.4 Current Gaps and the Rationale for the Present Study

Research highlights the potential of gamification to complements traditional teaching methods and making learning more interactive and enjoyable. Long-term studies are needed to examine how gamification affects both motivational and behavioral outcomes in specific fields like physics [17]. They highlight the need to examine both the motivational and behavioral aspects of gamification to fully understand its effectiveness.

In the field of physical research, gamified approaches can support the reproducibility of experiments and foster innovative teaching and learning methods. By integrating gamification elements into scientific experiments, students can be motivated to solve complex problems, develop creative solutions, and generate new insights. These gamified experiments can enhance collaboration, competition, and intrinsic motivation among participants, leading to more effective knowledge generation and dissemination. A study by Wang et al. [18] specifically examined the impact of gamification on teaching and learning in the context of the physical Internet, highlighting improvements in student satisfaction, knowledge acquisition, and test scores through gamified teaching methods.

Our review and analysis of the literature indicate, the integration of gamification and serious games into physics learning and research can improve the learning experience, increase motivation, and promote innovative teaching methods. By creating interactive and playful learning environments, both students and researchers can be encouraged to actively participate in the scientific process and deepen their understanding of physical principles. The research to date demonstrates the potential of gamification to transform education and establish sustainable learning practices that meet the needs of modern educational institutions.

## 2. Objectives and Hypotheses

### 2.1 Research Objectives

This study investigates the effects of gamification elements on student engagement, motivation, and learning outcomes within the context of secondary-level physics education. The primary objective of the Basketball Physics Challenge project is to examine how an interactive, game-based learning environment may enhance students' comprehension of fundamental physics concepts, including initial velocity, motion and trajectory, throwing angle, gravity, friction, and air resistance, by engaging them in an experiential, hands-on manner. A secondary objective was to explore structural performance modeling using Knowledge Space Theory (KST). This approach provides insights into latent learning trajectories, offering a deeper understanding of how gamification affects skill acquisition and progression beyond surface-level metrics. By modeling these trajectories, the study seeks to identify specific learning patterns that emerge in gamified and non-gamified conditions.

The study is guided by the following central research question: “How does the use of a gamified basketball application impact students’ motivation, engagement, and learning outcomes in the context of physics education?”

### 2.2 Hypotheses

In addressing this research question, the study proposes the following hypotheses:

**Hypothesis 1:** Gamification improves game performance**.** It is expected that participants in the gamification condition will exhibit higher engagement, evidenced by increased time on task, more attempts, and higher performance metrics (hits and scores) compared to those in the standard condition. This hypothesis will be tested by comparing average time on task, attempts, hits, and scores between the gamification and standard conditions using descriptive statistics and repeated measures ANOVA.

**Hypothesis 2:** Gamification enhances motivation. This hypothesis posits that participants in the gamification condition will report higher motivation scores compared to those in the standard condition. The hypothesis will be tested by analyzing self-reported motivation scores before and after each session using descriptive statistics and repeated measures ANOVA to compare changes in motivation across conditions and sessions.

**Hypothesis 3:** Gamification impacts quiz performance. It is hypothesized that participants in the gamification condition will achieve higher quiz scores compared to those in the standard condition. This hypothesis will be tested by comparing quiz scores before and after each session between the gamification and standard conditions using descriptive statistics and repeated measures ANOVA.

These hypotheses aim to provide deeper insights into the benefits of gamification and to identify which elements and strategies are most effective in increasing engagement, motivation, and learning efficiency in physics-based educational games. The study seeks to highlight the potential of gamification to sustain engagement and improve learning outcomes, with a focus on understanding its effects on knowledge retention and transfer.

## 3. Methods and Material

### 3.1 Participants

The study included 20 students from a secondary school in Liechtenstein who were in the optional 10$^{th}$ grade, an additional school year after completing the regular curriculum. The age range was between 15 and 18 years ($M = 16.25$, $SD = 0.76$), with 10 female and 9 male participants. 65% of the students were native speakers of German. Overall, the participants' physics scores tended to be in the lower range.

### 3.2 Study Design

In this study, a crossover design was used to investigate the effects of gamification on performance and skill acquisition in an educational game. The aim was to compare two versions of the game: one with gamification elements and one without. Participants were randomly divided into two groups, with one group starting with the gamification version and then switching to the non-gamification version after a break of several days, and the other group playing in the reverse order. This approach allowed individual differences among participants to be considered as a control variable while reducing the total number of participants required.

The gamification version of the game (gamification) included elements such as scoring systems, time limits, and competitions to increase participant motivation and engagement. The non-gamification version, referred to as version Standard, focused solely on the educational content without any additional game elements.

To measure physics comprehension, participants completed a physics comprehension quiz before playing their first game. After playing each game, both with and without gamification, they completed the same quiz again. The total score of the quiz was used as a measure of physics comprehension. The procedure was designed as follows: First, there was a pre-test before the first game, followed by a post-gamification test after playing the gamification version and a post-non-gamification test after playing the non-gamification version. This approach made it possible to isolate and measure the direct effects of gamification on learning progress. The consistent integration of the quiz into the different phases of the game ensured that the effects of gamification and non-gamification could be directly assessed on the participants' performance in relation to the physics concepts taught.

Additionally, a motivation questionnaire was administered at four key points to capture changes in participants' motivational levels. For both groups, the questionnaire was given before and after each game session. Participants who first played the gamification version completed the questionnaire before and after this session, and then again before and after playing the non-gamification version. Conversely, those who started with the non-gamification version followed the same procedure. This approach allowed for a detailed analysis of how gamification influenced motivation over time across different groups.


**Figure 1:** Screenshots of the Basketball Physics Challenge with the non-gamification version on the left and the gamification version on the right. This visual representation helps to illustrate the differences between the two game modes.

### 3.3 Game Design and Mechanics

The Basketball Physics Challenge is a web-based educational game designed to teach fundamental physics concepts such as trajectory, motion, and gravity. The game was developed using typical web technologies, i.e., HTML5, CSS, and JavaScript**,** to ensure accessibility across desktop and mobile platforms. Backend functionality (server-side scripting) is based on PHP; the log data are recorded in a mySQL database. For compliance reasons, the app does not use Cookies.

**Game Mechanics.** The mechanics were chosen to align with educational objectives, promoting active engagement and reinforcing physics concepts through interactive experimentation. Scoring and competition elements are grounded in motivational theory, emphasizing immediate feedback and goal-oriented learning. Key elements of the gamified version included:

**Scoring systems:** Players earn points based on their accuracy and the efficiency of their attempts.

**Time limits:** Levels must be completed within a set time, encouraging quick thinking and decision-making.

**Competition:** Progress trackers foster a sense of achievement and friendly competition.

The non-gamified version retained identical physics challenges without scoring, time constraints, or competitive elements, enabling a direct comparison of motivational influences.

To capture meaningful data, the game recorded player actions such as hits, misses, scores, and time spent on each level. These metrics provided insights into engagement and performance, forming the basis for subsequent analyses.

**Pilot Testing.** The app underwent a pilot phase to ensure usability and functionality. Feedback from testers informed refinements, such as simplified navigation, improved responsiveness for mobile devices, and precise simulation of physics principles, ensuring the game was intuitive and accessible for secondary school students.

### 3.4 Structural Performance Modeling

To achieve the study's objectives, we employed a combination of quantitative analyses and structural performance modeling. Specifically, Knowledge Space Theory (KST) was utilized to analyze skill acquisition patterns and learning trajectories among participants. This approach provides insights into the learning process beyond surface-level performance metrics, allowing for a deeper examination of how gamification influences the progression of physics understanding.

### 3.5 Level Structure

The structure of the game levels was designed to provide increasing complexity and challenge in relation to the physics concepts. For this structuring, a competency model was developed based on international curricula with a focus on Switzerland. The first levels focused on simple concepts such as initial velocity and throwing angle to teach players basic handling and understanding of trajectory. As the game progressed, the tasks became more complex and players had to consider advanced principles such as gravitational effects on different planets and the influence of environmental factors such as wind or vacuum. Each level aimed to teach a specific physics concept and deepen understanding through practical application.

### 3.6 Learning Task and Quiz Structure

The learning task focused on teaching and applying basic physics concepts in a simulated, interactive environment. This quiz was specifically designed to measure changes in understanding of physics principles such as initial velocity, motion, trajectory, throwing angle, gravity, friction, and air resistance. The practical application of these concepts took the form of a game in which players had to throw the ball into a basket with a cannon by adjusting these variables. This allowed participants to test hypotheses about physical effects and learn through direct feedback in the game. The overall quiz score was used as an indicator of physics understanding, with higher scores indicating deeper understanding.

### 3.7 Motivation Questionnaire

In addition to the physics comprehension quizzes, a motivation questionnaire was administered at four key points to capture changes in participants' motivational levels. The motivation scale was custom-built to assess constructs specific to physics gamification, including intrinsic/extrinsic motivation, interest in physics, satisfaction, and well-being. These constructs were selected based on their relevance to engagement in STEM education, as highlighted in prior studies [11, 16]. While standardized instruments exist, a tailored approach was necessary to capture domain-specific nuances in motivation and well-being. The administration and analysis of the questionnaire involved several steps. First, all variables in the questionnaire were identified and categorized into the following main categories: intrinsic motivation, extrinsic motivation, interest in physics, satisfaction, and well-being. To ensure uniform interpretation of the answer scales, responses to positively and negatively worded questions were recoded. Individual variables were then assigned to the five main categories based on their content. Intrinsic motivation included variables measuring internal incentives and engagement, while extrinsic motivation encompassed variables motivated by external rewards and recognition. Interest in physics measured participants' motivation and interest in physics concepts and the subject as a whole. Satisfaction assessed overall satisfaction with the game and its elements, and well-being measured emotional responses and general well-being during gameplay. For example, intrinsic motivation was assessed with items like, “I set personal goals when learning” and “I am motivated to achieve goals in digital games.” Extrinsic motivation included items such as "Rewards for learning achievements motivate me" and "I am motivated by the possibility to improve myself". Satisfaction included items such as, “I am satisfied with the challenges and difficulty level of my school tasks,” while interest in physics was gauged with statements like, “I am motivated for the subject of physics” and “I find physics interesting.” Well-being items measured emotional responses during gameplay, such as “I felt challenged” and “I felt happy. For each main category, an average score was calculated to represent the aggregated results of the associated variables.

By systematically categorizing and analyzing the responses, the study aimed to provide a comprehensive understanding of the motivational impact of gamification elements in the educational game.

### 3.8 Learning Outcome and Performance Measures

Learning progress was quantified using various performance measures, including average quiz scores before and after each game session. The overall quiz score was used as a measure of physics understanding. Descriptive statistics such as mean values and standard deviations were calculated to indicate learning progress and the effectiveness of gamification elements. Changes in quiz scores over time were examined to evaluate the impact of gamification on learning gains. A repeated measures analysis of variance (ANOVA) was used to compare differences between the gamification and non-gamification groups. This analysis allowed us to assess the direct effects of gamification on learning outcomes and performance. Furthermore, Pearson correlations between game performance metrics (time, attempts, hits, and scores) and quiz scores were calculated to explore the relationship between game performance and quiz performance. This helped to determine if gamification had a significant impact on knowledge acquisition and retention. Motivation was also assessed using a self-reported questionnaire administered before and after each game session. Descriptive statistics and repeated measures ANOVA were used to analyze changes in motivation scores across different conditions and sessions, providing insights into the motivational effects of gamification.

These measures provided a comprehensive overview of the participants' learning progress and helped to evaluate the effectiveness of the gamification elements in enhancing physics understanding and engagement. By systematically analyzing quiz performance, game metrics, and motivation scores, the study aimed to elucidate the potential benefits of gamification in educational games.

## 4. Results

The results of the Basketball Physics Challenge study provide detailed insights into the effectiveness of the game in teaching physics concepts and the role of gamification in learning. Initial analyses of group distribution across different measurement time points were visualized to illustrate player participation and engagement. Descriptive statistics, such as mean values and standard deviations, of the total scores for each group and measurement time point were calculated to indicate learning progress and the effectiveness of gamification elements.

### 4.1 Game Performance

In the analysis of the practice levels under different conditions, distinct patterns emerged between the gamification and standard approaches. Overall, we found slightly higher time on task and actions (e.g., number of attempts) as well as a slightly higher performance was observed in the practice levels in the gamification condition.

**Table 1.** Game performance in practice levels.

| | | Practice Level t1 (gamification first) | | | | Practice Level t2 (standard first) | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Time (sec) | Attempts | Hits | Score | Time (sec) | Attempts | Hits | Score |
| Gamification | **MEAN** | **161.64** | **29.64** | **18.45** | **266.36** | **122.88** | **22.63** | **12.88** | **192.50** |
| | SD | 105.40 | 19.41 | 16.97 | 169.01 | 72.28 | 18.81 | 17.29 | 207.83 |
| | MIN | 18.00 | 5.00 | 4.00 | 40.00 | 36.00 | 9.00 | 2.00 | 30.00 |
| | MAX | 287.00 | 74.00 | 60.00 | 600.00 | 266.00 | 67.00 | 54.00 | 650.00 |
| Standard | **MEAN** | **98.13** | **20.00** | **9.63** | **212.50** | **184.91** | **26.64** | **8.64** | **148.18** |
| | SD | 89.95 | 19.36 | 8.33 | 127.14 | 56.11 | 6.98 | 2.29 | 100.78 |
| | MIN | 6.00 | 6.00 | 2.00 | 40.00 | 99.00 | 13.00 | 5.00 | 50.00 |
| | MAX | 284.00 | 67.00 | 28.00 | 380.00 | 278.00 | 37.00 | 13.00 | 300.00 |

In the first session, the gamification condition resulted in higher time on task (M = 161.64 seconds, SD = 105.40), more attempts (M = 29.64, SD = 19.41), and higher performance metrics such as hits (M = 18.45, SD = 16.97) and scores (M = 266.36, SD = 169.01). In contrast, the standard condition showed lower values for time on task (M = 98.13 seconds, SD = 89.95), attempts (M = 20.00, SD = 19.36), hits (M = 9.63, SD = 8.33), and scores (M = 212.50, SD = 127.14).

In the second session, the gamification condition showed decreased performance with time on task (M = 122.88 seconds, SD = 72.28), attempts (M = 22.63, SD = 18.81), hits (M = 12.88, SD = 17.29), and scores (M = 192.50, SD = 207.83). Conversely, the standard condition exhibited a slight increase in performance with time on task (M = 184.91 seconds, SD = 56.11), attempts (M = 26.64, SD = 6.98), hits (M = 8.64, SD = 2.29), and scores (M = 148.18, SD = 100.78).

When analyzing game performance across different levels and sessions, there were also clear patterns between the gamification and standard approaches.

**Table 2.** Game performance in game levels.

| | | Game Levels t1 (gamification first) | | | | Game Levels t2 (standard first) | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Time (sec) | Attempts | Hits | Score | Time (sec) | Attempts | Hits | Score |
| Gamification | **MEAN** | **715.64** | **73.73** | **23.09** | **2346.36** | **554.25** | **64.38** | **19.88** | **2328.75** |
| | SD | 142.77 | 6.02 | 3.18 | 345.55 | 183.41 | 12.16 | 6.42 | 545.72 |
| | MIN | 562.00 | 61.00 | 18.00 | 1880.00 | 343.00 | 43.00 | 7.00 | 1230.00 |
| | MAX | 1074.00 | 81.00 | 28.00 | 2790.00 | 940.00 | 79.00 | 28.00 | 3100.00 |
| Standard | **MEAN** | **476.63** | **58.13** | **15.50** | **2085.00** | **584.55** | **67.18** | **24.73** | **2666.36** |
| | SD | 157.54 | 22.45 | 5.10 | 434.64 | 113.40 | 8.06 | 2.05 | 361.50 |
| | MIN | 250.00 | 21.00 | 9.00 | 1440.00 | 438.00 | 54.00 | 21.00 | 2100.00 |
| | MAX | 697.00 | 85.00 | 23.00 | 2790.00 | 940.00 | 79.00 | 28.00 | 3100.00 |

In the first session, the gamification condition resulted in higher time on task (M = 715.64 seconds, SD = 142.77), more attempts (M = 73.73, SD = 6.02), and higher performance metrics such as hits (M = 23.09, SD = 3.18) and scores (M = 2346.36, SD = 345.55). In contrast, the standard condition showed lower values for time on task (M = 476.63 seconds, SD = 157.54), attempts (M = 58.13, SD = 22.45), hits (M = 15.50, SD = 5.10), and scores (M = 2085.00, SD = 434.64).

In the second session, about a week later, the gamification condition showed decreased performance with time on task (M = 554.25 seconds, SD = 183.41), attempts (M = 64.38, SD = 12.16), hits (M = 19.88, SD = 6.42), and scores (M = 2328.75, SD = 545.72). Conversely, the standard condition exhibited an increase in performance with time on task (M = 584.55

seconds, SD = 113.40), attempts (M = 67.18, SD = 8.06), hits (M = 24.73, SD = 2.05), and scores (M = 2666.36, SD = 361.50).

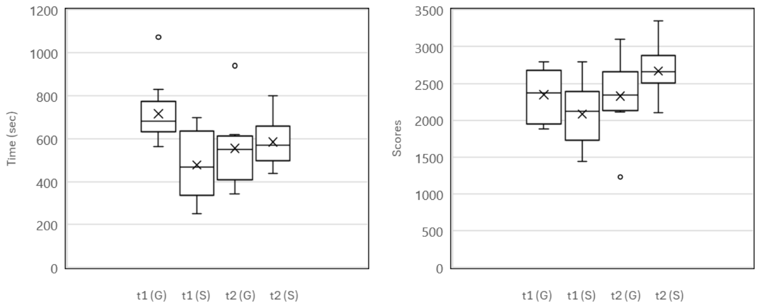


**Figure 2.** Box plots for time on task (left) and scores (right) across sessions and conditions. t1 and t2 represent the first and second sessions, respectively, while G and S denote gamification and standard conditions. The plots illustrate variations in engagement (time) and performance (scores) between conditions and sessions

Figure 2 shows the box plots of time on task as well as scores. The group that performed the gamification condition first showed a clear decrease in time on task when performing the standard condition about a week later; whereas the group that performed the standard condition first, showed a clear increase in time on task. This is a clear indication that the gamification elements increased the efforts of playing the game. With respect to a possible learning gain, it is reasonable to assume that students in both conditions encounter an increase in their skills.

A repeated measures analysis of variance (ANOVA) yielded significant effects for the dependent variable time on task for condition (gamification vs standard; $F(1, 17) = 4.455$, $p = .50$, partial $\eta^2 = .208$) and for the interaction of session * condition ($F(1, 17) = 17.025$, $p < .001$, partial $\eta^2 = .50$). For hits, the repeated measures ANOVA revealed significant effects for session ($F(1, 17) = 9.049$, $p = .008$, partial $\eta^2 = .347$) and condition ($F(1, 17) = 13.345$, $p < .002$, partial $\eta^2 = .44$), with no significant interaction found. For scores, the repeated measures ANOVA indicated significant effects for session ($F(1, 17) = 5.623$, $p = .030$, partial $\eta^2 = .249$) and a marginal effect for condition ($F(1, 17) = 3.835$, $p = .067$, partial $\eta^2 = .184$), with no significant interaction observed.

### 4.2 Quiz Performance

To provide insights into the effectiveness of the different conditions, quiz performance was analyzed before and after each session.

**Table 3.** Quiz performance in pretest and posttests after each session.

| | | Pre | Post t1 | Post t2 |
|---|---|---|---|---|
| Gamification | **MEAN** | **8.91** | **8.18** | **6.63** |
| | SD | 2.55 | 3.19 | 2.56 |
| | MIN | 4.00 | 1.00 | 3.00 |
| | MAX | 13.00 | 11.00 | 10.00 |
| Standard | **MEAN** | **8.50** | **6.63** | **7.82** |
| | SD | 2.27 | 2.13 | 2.18 |
| | MIN | 5.00 | 4.00 | 6.00 |
| | MAX | 12.00 | 9.00 | 12.00 |

The quiz results in the pretest before the first session did not differ substantially between conditions, with the gamification condition showing a mean score of M=8.91, SD = 2.55, while the standard condition had a mean score of M=8.50, SD = 2.27. After the first session, the quiz scores were higher for the gamification condition (M=8.18, SD = 3.19) compared to the standard condition (M=6.63, SD = 2.13). However, after the second session, the quiz results were opposite, with the gamification condition scoring lower (M=6.63, SD = 2.56) and the standard condition scoring higher (M=7.82, SD = 2.18).

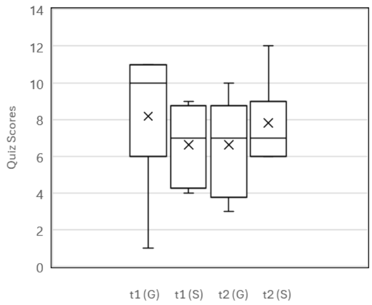


**Figure 3.** Quiz performance in the posttests after each session. t1 and t2 represent the first and second sessions, respectively, while G and S denote the gamification and standard conditions (e.g., t2(G) indicates results for the gamification condition in the second session).

Figure 3 summarizes the results as box plots. A repeated measures ANOVA did not result in significant differences in the test scores between conditions and sessions. This indicates that there is only a moderate transfer between skills gained in the game, independent of condition, to the knowledge quiz. The decrease from the quiz prior to using the learning app can be explained with a general loss in motivation.

**Table 4.** Correlations between Game performance and quiz points.

| | | Session 1 | | | Session 2 | | |
|---|---|---|---|---|---|---|---|
| | | Time | Score | Quiz | Time | Score | Quiz |
| Gamification | Time | 1 | -0.426 | -0.018 | 1 | -.805* | -0.26 |
| | Score | -0.426 | 1 | 0.257 | -.805* | 1 | 0.675 |
| | Quiz | -0.018 | 0.257 | 1 | -0.26 | 0.675 | 1 |
| Standard | Time | 1 | -0.296 | -0.026 | 1 | -0.137 | -0.175 |
| | Score | -0.296 | 1 | 0.306 | -0.137 | 1 | 0.329 |
| | Quiz | -0.026 | 0.306 | 1 | -0.175 | 0.329 | 1 |

*Statistically significant at the $p < 0.05$ level.

The Pearson correlations between game performance (time and scores) and quiz points were rather low in the first session (cf. Table 4). In the second session, however, we found a strong positive correlation between game scores and quiz points in the gamification condition ($r = 0.675$) and moderate positive correlation in the standard condition ($r = 0.329$). This might indicate that effects on the knowledge test are only achieved after using the app twice. Interestingly, there were negative correlations between time on task and game scores throughout, suggesting that skilled learners needed less time to perform the game tasks. There were no correlations between time on task and quiz points.

### 4.3 Motivation

To evaluate the impact of different conditions on student motivation, self-reported motivation scores were analyzed before and after each session.

**Table 5.** Self-reported motivation scores.

| | | Session 1 | | Session 2 | |
|---|---|---|---|---|---|
| | | Pre | Post | Pre | Post |
| Gamification | **MEAN** | **3.01** | **3.02** | **2.93** | **2.68** |
| | SD | 0.20 | 0.26 | 0.42 | 0.39 |
| | MIN | 2.82 | 2.56 | 2.08 | 2.11 |
| | MAX | 3.46 | 3.58 | 3.63 | 3.17 |
| Standard | **MEAN** | **2.94** | **2.15** | **3.12** | **2.84** |
| | SD | 0.29 | 0.67 | 0.25 | 0.46 |
| | MIN | 2.41 | 1.00 | 2.86 | 1.75 |
| | MAX | 3.24 | 3.07 | 3.71 | 3.45 |

Overall, motivation scores were in a medium range throughout. For the gamification condition, the motivation scores before the first session (M = 3.01, SD = 0.20) were similar to those after the first session (M = 3.02, SD = 0.26). In the second session, the motivation scores before playing the non-gamification version were slightly lower (M = 2.93, SD = 0.42), and they decreased further after the session (M = 2.68, SD = 0.39).

For the standard condition, the motivation scores before the first session (M = 2.94, SD = 0.29) were higher than those after the first session (M = 2.15, SD = 0.67). In the second session, the motivation scores before playing the gamification version were higher (M = 3.12, SD = 0.25), and they decreased slightly after the session (M = 2.84, SD = 0.46).

Overall, motivation prior to using the app did not change significantly from session 1 to session 2. However, a distinct effect of the condition was observed for motivation after using the app. There was a noticeable decrease in motivation when participants first experienced the gamification condition and then, about a week later, the standard condition. The average decrease was -0.18. Conversely, when participants first performed the standard condition and then the gamification condition, the post-motivation score increased by 0.53.

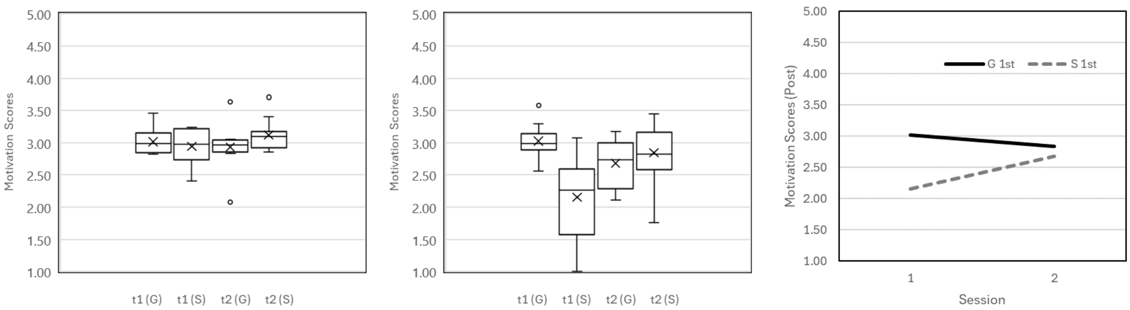


**Figure 4.** Motivation scores for both sessions (t1, t2) and both conditions (gamification, standard). The left diagram shows the motivation scores before using the app, the middle diagram that after using the app. The right diagram shows the interaction effect of condition and session.

A repeated measures analysis of variance (ANOVA) yielded significant effects for condition (gamification vs standard; $F(1, 17) = 5.491$, $p = .032$, partial $\eta^2 = .244$), whereas the linear model did not reveal significant differences for measurement time (prior and after session1 and session 2) and no significant interaction of condition and measurement time. When looking at the self-reported motivation after both sessions, the repeated measures ANOVA yielded statistically significant differences for condition ($F(1, 17) = 9.428$ $p = .007$, partial $\eta^2 = .357$)

and a significant interaction of condition * session F(1, 17) = 8.023, p = .011, partial $\eta^2$ = .321). This is a clear indication for the strong motivational potential of the gamification elements independent of a repeated use of the learning app (cf. Figure 4, right diagram).

These results indicate that the motivational effects of gamification are more complex than initially anticipated and may depend on the order in which the conditions are experienced. The gamification elements can significantly impact motivation, but this effect may vary based on the participants' prior experiences with the game.

### 4.4 Correlations Among Motivation, Quiz Performance, and Game Metrics

To understand the relationships among motivation, quiz performance, and game metrics, a correlation analysis was conducted.

**Table 6.** Correlation matrix for motivation, quiz points, and game performance.

| | MOT_1_pre | MOT_1_post | MOT_2_pre | MOT_2_post | Q1 | Q2 | Q3 | Time_1 | Score_1 | Time_2 | Score_2 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| MOT_1_pre | 1 | .594** | .666** | 0.424 | 0.449 | 0.399 | 0.138 | -0.183 | .591** | -.592** | .644** |
| MOT_1_post | .594** | 1 | .641** | 0.336 | 0.368 | 0.226 | 0.305 | 0.329 | .569* | -0.378 | .671** |
| MOT_2_pre | .666** | .641** | 1 | .616** | 0.039 | 0.1 | 0.262 | 0.006 | 0.309 | -0.364 | .697** |
| MOT_2_post | 0.424 | 0.336 | .616** | 1 | 0.017 | 0.141 | 0.336 | 0.05 | 0.08 | -0.082 | 0.4 |
| Q1 | 0.449 | 0.368 | 0.039 | 0.017 | 1 | .673** | .571* | 0.133 | 0.322 | -0.446 | .519* |
| Q2 | 0.399 | 0.226 | 0.1 | 0.141 | .673** | 1 | .630** | 0.164 | 0.331 | -0.247 | .571* |
| Q3 | 0.138 | 0.305 | 0.262 | 0.336 | .571* | .630** | 1 | 0.279 | 0.123 | -0.184 | .558* |
| Time_1 | -0.183 | 0.329 | 0.006 | 0.05 | 0.133 | 0.164 | 0.279 | 1 | -0.047 | .624** | 0.059 |
| Score_1 | .591** | .569* | 0.309 | 0.08 | 0.322 | 0.331 | 0.123 | -0.047 | 1 | -.586** | 0.341 |
| Time_2 | -.592** | -0.378 | -0.364 | -0.082 | -0.446 | -0.247 | -0.184 | .624** | -.586** | 1 | -.477* |
| Score_2 | .644** | .671** | .697** | 0.4 | .519* | .571* | .558* | 0.059 | 0.341 | -.477* | 1 |

* Statistically significant on the 5% level

** Statistically significant on the 1% level

Table 6 summarizes the Pearson correlations among the variables, motivation, quiz points, and game performance across two sessions. Overall, we found strong positive correlations between the self-reported motivation at all measurement times, with coefficients of 0.594**, 0.666**, and 0.424 for MOT_1_post, MOT_2_pre, and MOT_2_post (prior and after each session). This may be considered an indicator for the stability of self-reported motivation values.

Motivation was found to be positively correlated with quiz results, although this relationship was only moderate and not systematic across all variables. We found moderate correlations between the motivation prior to using the app and the quiz after using the app in both conditions, particularly with coefficients of 0.449 for Q1 and 0.399 for Q2 (indicated by light blue cells in Table x). This suggests that participants who start with higher motivation generally perform better on the quizzes, although this relationship is not uniformly strong across all quiz points.

More importantly, the motivation prior to using the app showed a strong correlation with the game scores (Score_1: r = 0.591**, Score_2: r = 0.644**), suggesting that participants who start highly motivated achieve higher scores in the game. (indicated by light orange cells in Table x).

Additionally, the motivation after using the app was positively correlated to scores (MOT_1_post and Score_1: r = 0.569*; MOT_2_post and Score_2: r = 0.4. There were no or negative correlations between motivation and time on task. As mentioned before, there appears being a negative correlation between time on task and scores.

Finally, we found a strong positive correlation between time on task in sessions 1 and 2 (Time_1 and Time_2: r = 0.624**) (indicated by the pink cell in Table 6). This might indicate that the time spent on tasks is independent of the motivational level, while diligence (i.e., scores) isn't.

### 4.5 Condition-Specific Correlations Among Motivation, Quiz Performance, and Game Metrics

To delve deeper into the impact of different conditions on motivation, quiz performance, and game metrics, we conducted a condition-specific correlation analysis across both sessions.

**Table 7.** Correlation matrix by condition and session.

| | | Session 1 | | | | | Session 2 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Mot_pre | Mot_post | Q_post | Time | Score | Mot_pre | Mot_post | Q_post | Time | Score |
| Gamification | Mot_pre | 1 | 0.192 | 0.398 | -0.179 | 0.187 | 1 | 0.623 | 0.636 | -.805* | .942** |
| | Mot_post | 0.192 | 1 | -0.083 | 0.013 | -0.231 | 0.623 | 1 | 0.317 | -0.496 | 0.701 |
| | Q_post | 0.398 | -0.083 | 1 | -0.018 | 0.257 | 0.636 | 0.317 | 1 | -0.26 | 0.675 |
| | Time | -0.179 | 0.013 | -0.018 | 1 | -0.426 | -.805* | -0.496 | -0.26 | 1 | -.805* |
| | Score | 0.187 | -0.231 | 0.257 | -0.426 | 1 | .942** | 0.701 | 0.675 | -.805* | 1 |
| Standard | Mot_pre | 1 | .894** | 0.415 | -0.55 | .880** | 1 | .639* | -0.374 | 0.324 | 0.184 |
| | Mot_post | .894** | 1 | 0.177 | -0.355 | .860** | .639* | 1 | 0.3 | 0.258 | 0.087 |
| | Q_post | 0.415 | 0.177 | 1 | -0.026 | 0.306 | -0.374 | 0.3 | 1 | -0.175 | 0.329 |
| | Time | -0.55 | -0.355 | -0.026 | 1 | -0.296 | 0.324 | 0.258 | -0.175 | 1 | -0.137 |
| | Score | .880** | .860** | 0.306 | -0.296 | 1 | 0.184 | 0.087 | 0.329 | -0.137 | 1 |

* Statistically significant on the5% level
** Statistically significant on the 1% level

Table 7 summarizes the Pearson correlations among the variables, separated by condition (gamification, standard) and sessions (1 and 2). In the gamification condition during Session 1, initial motivation showed a weak correlation with subsequent motivation (r = 0.192), indicating that initial motivation had a minimal influence on subsequent motivation. Additionally, initial motivation was moderately associated with better quiz performance (r = 0.398), suggesting that higher initial motivation was moderately associated with better quiz results. Interestingly, a moderate negative correlation was observed between time on task and game scores (r = -0.426), indicating that more efficient players, who achieved higher scores, spent less time on the task.

In Session 2, the correlations in the gamification condition showed notable changes. Initial motivation had a strong influence on subsequent motivation (r = 0.623) and quiz performance (r = 0.636), suggesting that initial motivation had a significant influence on subsequent motivation after the second session and also a sustained link between higher initial motivation and better quiz results. A very strong correlation was found between initial motivation and game scores (r = 0.942**), suggesting that initial motivation significantly influenced game performance. The negative correlation between time on task and game scores was also strong (r = -0.805*), indicating that higher scores were associated with shorter gameplay duration, reinforcing the notion that more skilled players completed tasks more efficiently.

In the standard condition during Session 1, initial motivation was strongly correlated with subsequent motivation (r = 0.894**) and game scores (r = 0.880**), indicating high stability of motivation across the session and suggesting that high initial motivation significantly improved game performance. A moderate correlation existed between initial motivation and quiz performance (r = 0.415). The correlation between time on task and game scores was moderately negative (r = -0.296).

In Session 2 of the standard condition, the correlation between initial motivation and subsequent motivation remained strong (r = 0.639*), but the correlation with quiz performance turned negative (r = -0.374), suggesting that higher initial motivation did not necessarily lead to better quiz results in the second session. The correlation with game scores weakened (r = 0.184), and the negative correlation between time on task and game scores was minimal (r = -0.137).

Overall, the analysis indicates that pre-motivation plays a more critical role in performance in the standard condition compared to the gamification condition. In the standard condition, high initial motivation was strongly linked to better game performance and stable motivation levels across sessions. Conversely, in the gamification condition, the impact of initial motivation on game performance and subsequent motivation was less pronounced in the first session but became more significant in the second session. This suggests that gamification elements may help sustain motivation and improve performance over time, but initial motivation plays a crucial role in non-gamified settings. These findings highlight the importance of initial motivation in enhancing performance, particularly in environments without gamification elements.

### 4.6 Structural Performance Modelling.

Beyond, the summative performance values, such as scores, we are interested in the structure of performance, i.e., the relationship of levels in the app and their underlying latent skills. This understanding serves (a) a qualitative and formative diagnosis of skills/knowledge and (b) the design of the app (e.g., level sequencing). To elucidate the performance structure, we cannot only refer to the solution frequencies of levels, we need to consider the participants' patterns of correctly mastered levels. Also, we need to account for errors (slips) and luckily mastering a level (guessing). This is of particular importance in environment with large degrees of freedom, as it is the case in the presented (gamified) learning app.

Methodologically, the structural analyses are based on Knowledge Space Theory (KST; [19], [20], [ 21]. KST is a combinatorial approach from the family of Cognitive Diagnostic Models (CDM; [22], which postulates a set of items (in our case levels), which are in a non-linear relation, a so-called surmise relation. From mastering one item/level, one can surmise the mastery of one or more other items/levels. For the design of the nine learning levels, we decomposed the domain into 11 skills from the domain of ballistic trajectories. On this basis we designed the nine levels (Table 8).

**Table 8.** Assignment of skills to levels.

| | L1 | L2 | L3 | L4 | L5 | L6 | L7 | L8 | L9 |
|---|---|---|---|---|---|---|---|---|---|
| Acceleration | x | x | x | x | x | x | x | x | x |
| Gravity (earth) | x | x | x | x | x | x | x | x | x |
| Air resistance | x | x | x | x | x | | | | x |
| Shooting angle | | x | x | x | x | x | x | x | x |
| Headwind, tailwind | | | x | | | | | | x |
| Shooting height | | | | x | | | | | x |
| Magnet affecting balls | | | | | x | | | | |
| Gravity (moon) | | | | | | | x | | |
| Buoyancy in water | | | | | | | | x | |
| Vacuum | | | | | | x | | | |
| Least Acceleration | | | | | | | | | x |

The skills are not equally difficult and are typically acquired/taught along a certain learning trajectory [23] This implies a non-linear competence structure, which, in turn, implies a “difficulty” structure among the levels, the knowledge space. This includes the empty set and the full set (powerset). Depending on the hypothesized learning trajectory for the skills, different knowledge space models can be obtained. Figure 5 gives an overview of the three models, we compared in this study. Model 1 is a linear model based on a linear sequential increase of level difficulty. This model was the basis for the level design. Model 2 is an alteration of Model 1, based on the observed solution frequencies in the data. Model 3 is a non-linear model based on a theoretical learning trajectory for this domain (cf., [24]).

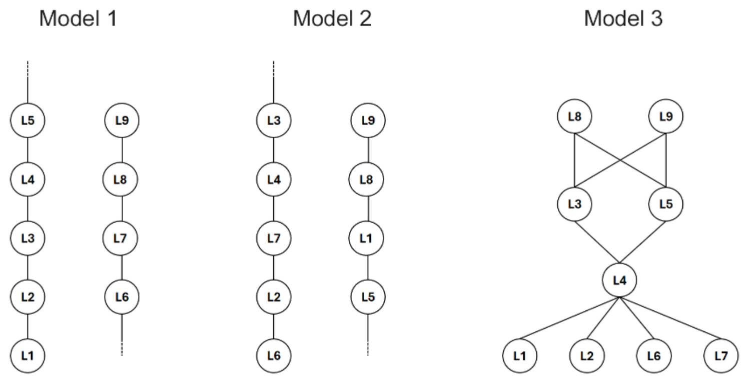


**Figure 5.** Structural performance models investigated in this study.

Based on a model, KST allows deriving a knowledge space, which is the collection of all admissible knowledge states of a learner[25]. For example, assuming that Model 1 is correct, a particular learner may master Level 1 and Level 2, however, it would not be admissible according to the model, that a learner masters Level 2 but not Level 1. Of course, the observable behavior is prone to careless errors (slips) or lucky guesses. Therefore, a probabilistic analyses of answer patterns (i.e., a learner’s set of mastered levels) is necessary. CDM in general and KST in particular offer a range of methods for model validation. These methods investigate the consistency of empirical answer patterns and the patterns hypothesized by a knowledge (or competence) structure. Different numerical measures have been proposed, for example, the Discrepancy Index [26] or the Distance Agreement Coefficient [27] or the Minimal Symmetric Set Difference [28], which likely is the most common metric. We applied the method of the minimal symmetric set distance (MSSD; [28]), which computes the distance of an answer pattern to the nearest knowledge state (cf. [29]). For example, for an answer pattern {1, 0, 1} (where 1 denotes the mastery of an item/level) the knowledge state {1, 1, 1} has a set distance of 1. For a model of nine levels, the minimum MSSD is 0 and the maximum 4. According to the KST logic, the minimum number of knowledge states in a model is 2 and the maximum is 2N, in our case 29 = 512.

First, we derived the knowledge spaces for the three models (Figure 5). Second, we coded the raw log data into mastered and non-mastered levels for each participant. Given that each of the levels requires shooting four different balls into the basket, there is a certain degree of freedom in coding these actions dichotomously. Therefore, we once coded three or four hits per level (i.e., hitting the basket with all four balls) as mastered (Table 9, “by 3”) and once only by four hits (Table 9, “by 4”). Third, we computed the average MSSD (cf. Table 9). For both coding variants, Model 2 yielded the least MSSD. However, in model fitting, there is always a size-fit tradeoff. This means that the larger a model is, the better is its fit only by chance. With a model the includes all 512 states, a perfect fit would be obtained to any possible data set, however, there is no explanatory value to such model.

**Table 9.** Results of the KST-based structural analyses.

| | | Cardinality | MSSD | SD | z-value | p-value | SFTOC |
|---|---|---|---|---|---|---|---|
| by 3 | Model 1 | 10 | 1.2000 | 0.9677 | 6.9188 | < .001 | 0.0279 |
| | Model 2 | 11 | 0.8444 | 0.9034 | 7.6740 | < .001 | 0.0272 |
| | Model 3 | 7 | 1.1778 | 0.9603 | 6.8047 | < .001 | 0.0194 |
| by 4 | Model 1 | 10 | 1.4000 | 1.0090 | 6.9188 | < .001 | 0.0300 |
| | Model 2 | 11 | 1.0000 | 0.7071 | 7.2852 | < .001 | 0.0286 |
| | Model 3 | 7 | 1.1778 | 0.7772 | 6.3047 | < .001 | 0.0194 |

As shown in Figure 6 (left panel), the size-fit ratio can be displayed as the percentages of a models fit and size (cardinality) to the maximum fit and size (which is 100%; in our case the maximum fit equals the minimum distance of zero and the maximum size is 512 states). The diagonal of the diagram indicates an increase of model fit with increasing size solely due to chance. The more a model is located above the diagonal, the higher is its explanatory value (the approach is described in more detail in Albert, Kickmeier-Rust, & Matsuda [29]. As shown in Figure 6, this is the case for all models and coding variants. For a statistical assessment, we can compute the probability of the standardized distances of model and diagonal, using a Gauss test). As shown in Table 9, all models have statistically significant distances to the diagonal, meaning that their explanatory value is significantly above chance. A further model selection criterion is the size-fit tradeoff coefficient (i.e., rel. size / rel. fit). In conclusion, we suggest Model 3 being the most explanatory; it has the second-best fit (MSSD) and the lowest SFTOC. Furthermore, the coding variant "by 3" (three or four hits) leads to a slightly better model fit.

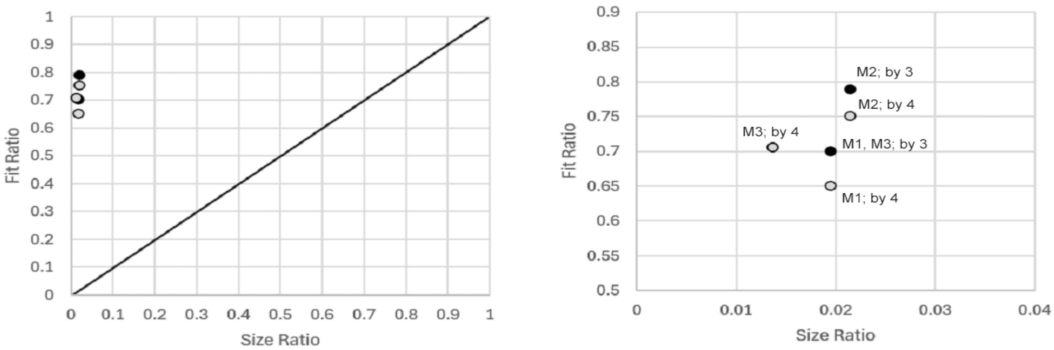


**Figure 6.** MSSD-based size-fit tradeoff. The left panel shows the full size-fit scale; the right panel shows a cutout to enable a better location of the models' values.

Based on this conclusion, one could draw conclusions on the nature of the learning trajectory in this domain (cf. Table 10). Figure 7 illustrates a likely learning trajectory based on Model 3. This model, of course, is limited by the rather rough granularity of the levels as diagnostic entity.

**Table 10.** Assignment of skills (S) to Knowledge States (K) of Model 3.

| | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 | S9 | S10 | S11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **K1** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **K2** | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 |
| **K3** | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 |
| **K4** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 |
| **K5** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| **K6** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 |
| **K7** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

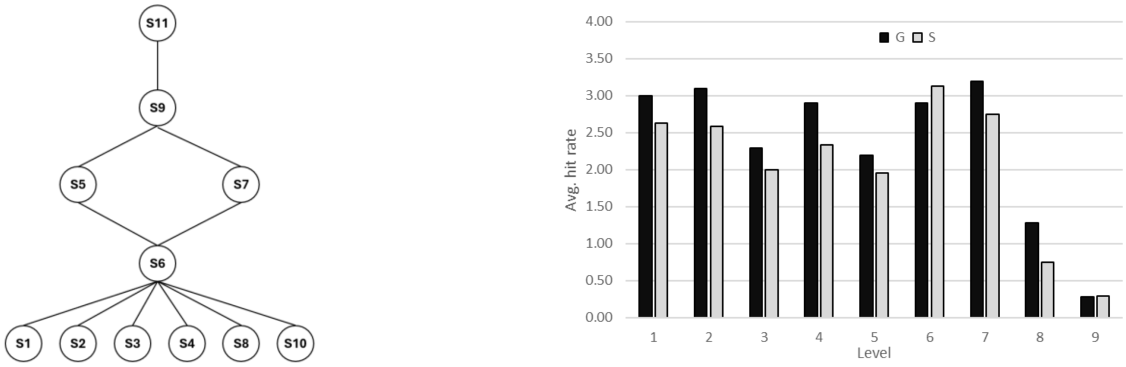

**Figure 7.** Structural representation of the learning trajectory within the specified domain (left panel). Comparative analysis of average hit rates (number of successful basket hits) observed in the gamification (G) versus the standard (S) conditions across nine distinct levels (right panel).

The models do not appear affected by the conditions (gamification vs standard). As shown in Figure 7 (right panel), the average hit rates do not differ substantially between conditions – with the exception of the moderate advantage of the gamification condition – as reported in section "game performance".

## 5. Discussion

The primary objective of this study was to examine the effects of gamification on learning performance, motivation, and quiz outcomes in a basketball-themed educational game designed to teach basic physics concepts. Our findings provide comprehensive insights into how gamification elements, such as points, time limits, and rewards, influence various aspects of learning and engagement.

**Impact of Gamification on Game Performance.** The results indicate that gamification significantly enhanced game performance. Participants in the gamification condition demonstrated higher engagement, evidenced by increased time on task and more attempts. This aligns with Hypothesis 1, which posited that gamification would improve game performance. Specifically, the gamification condition led to higher performance metrics, including hits and scores, compared to the standard condition. These findings are consistent with previous research suggesting that gamification can enhance engagement and motivation, leading to improved performance in educational contexts [30], [31], [32].

Interestingly, while the gamification condition initially resulted in better performance, this advantage diminished in the second session. This could be attributed to motivational or fatigue effects, suggesting that the novelty of gamification may wear off over time, impacting sustained engagement. Mitchell et al. [17] emphasize the importance of long-term studies to understand how gamification affects behavior and outcomes, highlighting the need to examine both the motivational and behavioral aspects to fully understand its effectiveness. Future research should explore strategies to maintain the motivational benefits of gamification over extended periods.

**Influence on Quiz Performance.** The study also investigated the impact of gamification on quiz performance, aiming to understand the transfer of skills from the game to knowledge assessments. The results showed a moderate improvement in quiz performance after the first session in the gamification condition. However, this effect did not persist in the second session, and overall quiz results showed no significant differences between conditions. This finding partially supports Hypothesis 3, indicating that while gamification can enhance immediate learning outcomes, its impact on knowledge retention and transfer is limited [15], [16]. The observed correlation between game scores and quiz performance in the second session suggests that gamification may facilitate skill consolidation over time. This finding aligns with prior research emphasizing the delayed benefits of interactive learning tools, where repeated engagement strengthens knowledge retention and transfer [15, 33]. However, the lack of consistent improvement across sessions indicates that while gamification can enhance immediate engagement, its effects on long-term retention require further exploration.

The Pearson correlations between game performance and quiz points were rather low in the first session but showed a strong positive correlation between game scores and quiz points in the gamification condition in the second session. This suggests that repeated use of the gamified learning app may strengthen the relationship between game performance and knowledge acquisition [33], [34].

**Effects on Motivation.** The motivational impact of gamification was assessed through self-reported motivation scores. The results revealed that motivation remained stable across sessions for participants who experienced the gamification condition first. However, a notable decrease in motivation was observed for participants who switched from the gamification condition to the standard condition. Conversely, motivation increased for participants who started with the standard condition and then experienced the gamification condition. These findings support Hypothesis 2, suggesting that gamification can enhance motivation, but this effect is influenced by the order of game types [35], [36].

The repeated measures ANOVA highlighted significant effects for condition and a significant interaction of condition and session, indicating the strong motivational potential of gamification elements independent of repeated use of the learning app. These results align with existing literature on the motivational benefits of gamification [14], [16].

The motivational decline observed in the standard condition underscores the importance of sustaining engagement through gamification elements. However, variability in quiz performance across sessions suggests that the design of game mechanics requires further refinement to align with learning objectives more closely. For instance, integrating adaptive elements that adjust difficulty based on player performance might help maintain engagement and support learning outcomes.

### 5.1 Comparison with Previous Research

Research has shown that gamification can have a positive impact on short- and long-term knowledge gain, suggesting that it has the potential to improve students' understanding of challenging topics such as physics [34]. Comparative studies have been conducted to assess the differences between gamified and non-gamified learning environments[34]. By incorporating gamification elements, teachers can create more engaging and interactive learning experiences that help students better understand difficult topics and improve their long-term academic performance [33].

Gamification has been associated with improved learning performance and intrinsic motivation in education, suggesting that it has the potential to enhance students' understanding and mastery of specific academic content [16]. The integration of SRL strategies within gamification can support deeper learning and the application of learning to new tasks. Furthermore, gamified learning environments have been shown to promote active participation and proactivity among students, leading to better learning outcomes [37].

### 5.2 Structural Performance Modelling

To further understand the structure of performance, we employed Knowledge Space Theory (KST) to analyze the relationships between levels and underlying latent skills. Using KST, we identified key learning trajectories and skill acquisition patterns, which suggest that gamified environments facilitate complex and non-linear learning processes. The structural analyses revealed that Model 2, based on observed solution frequencies, yielded the least Minimal Symmetric Set Distance (MSSD), suggesting it was the best fit for our data. This model provided valuable insights into the learning trajectory and skill acquisition patterns, highlighting the complexity and non-linearity of learning in a gamified environment [38].

The findings from SPM underscore the importance of designing game levels that align with learners' skill progression and knowledge acquisition patterns. The identification of optimal learning trajectories through Model 3 highlights the value of non-linear learning paths in gamified environments. For game designers, these insights suggest the need to create adaptive level sequences that cater to diverse learner abilities. For instance, incorporating branching pathways or scaffolding mechanisms can allow players to progress at their own pace while addressing knowledge gaps.

Moreover, the ability to identify slips and guesses in level completion offers an opportunity for real-time feedback systems. By integrating diagnostic algorithms informed by SPM, educational games can dynamically adjust their difficulty, fostering a balance between challenge and mastery. This adaptive approach could enhance learner engagement while ensuring that educational objectives are met. The application of KST in this study underscores the importance of considering individual learning paths and the varied progression of skills in educational game design.

## 6. Limitations and Future Research

While this study provides valuable insights into the effects of gamification on learning and motivation, it has several limitations. The sample size was relatively small (N=19) and from a single secondary school, which may limit the generalizability of the findings. A larger sample size from multiple schools or educational settings would provide a more robust basis for generalizing the results.

Given the small sample size and the use of parametric tests like ANOVA may have limitations. Non-parametric analyses, such as the Wilcoxon signed-rank test, were considered as alternatives that do not assume normal distribution of data. However, to maintain consistency with existing research that often employs ANOVA for similar studies, we proceeded with parametric tests. Future studies should include larger and more diverse samples and consider employing non-parametric tests when appropriate, especially if assumptions of normality are not met.

Additionally, the study focused on a specific educational game and set of physics concepts, which may not be representative of other subjects or gamified learning environments. Future research should explore larger and more diverse samples to validate the findings and extend the investigation to other educational contexts.

Longitudinal studies are also needed to examine the long-term effects of gamification on learning outcomes and motivation [17], [32], [36] and also to assess whether gamification fosters sustained interest in physics and other STEM disciplines over time

Another promising avenue for future research is the examination of the effects of personalized gamification. Adaptive learning technologies that tailor game elements to individual student needs and preferences could enhance the effectiveness of gamification. Studies could investigate how personalized gamified experiences impact learning outcomes compared to a one-size-fits-all approach.

Additionally, exploring the role of teacher involvement in gamified learning environments could yield important insights. Teachers' attitudes towards gamification, their training in implementing gamified strategies, and their interactions with students in a gamified setting are critical factors that could influence the success of gamified learning interventions. Research in this area could help develop best practices for integrating gamification into classroom instruction.

### 6.1 Implications for Educational Practice.

The findings of this study have important implications for educational practice. Gamification appears to be a promising approach to enhance engagement and performance in educational games. Educators and game designers should consider incorporating game elements strategically to sustain motivation and improve learning outcomes. However, it is crucial to balance gamification elements to avoid potential fatigue effects and ensure that the novelty of the game does not diminish over time [35].

Moreover, the integration of structural performance modelling techniques, such as KST, can provide a deeper understanding of learners' skill acquisition and inform the design of more effective and personalized learning experiences [25].Studies by Hamza and Tóvölgyi [38] indicate that gamified e-learning platforms can improve learner engagement and knowledge retention, suggesting that similar approaches could be successfully applied in education to improve understanding of complex concepts and increase student engagement.

## 7. Conclusions

This study aimed to answer the overarching research question: How does gamification impact engagement, motivation, and learning outcomes in secondary physics education, particularly in a basketball-themed game app? This study demonstrates that gamification can significantly enhance student engagement and performance in physics education, with a positive but moderate impact on quiz performance. The research focused on a basketball-themed educational game designed to teach fundamental physics concepts, such as initial velocity, motion, and trajectory. The motivational benefits of gamification were evident, although they varied depending on the sequence in which different game types were experienced. By employing structural performance modeling based on Knowledge Space Theory (KST), we gained valuable insights into learning trajectories and skill acquisition patterns in a gamified environment. These insights are particularly relevant for physics education, where understanding complex and abstract concepts is crucial.

Our findings contribute to the growing body of evidence supporting the use of gamification in educational settings, suggesting it can foster deeper engagement and improved learning outcomes in physics. However, the study also highlights the need for further research to explore the long-term effects of gamification and its broader applicability across different educational contexts. This research underscores the potential of gamification to transform educational practices in physics by making learning more interactive and enjoyable. It also emphasizes the importance of strategic implementation to sustain the benefits of gamification over time, ensuring that the novelty and motivational aspects do not diminish.

## References


[1] K. Zuber, “Die Entdeckung der solaren Neutrinos,” Physik in Unserer Zeit, 53(3), 150-150, 2022. https://doi.org/10.1002/piuz.202270309

[2] L. Schwarz, S. Kaiser, & D. Manske, “Higgs-Bosonen im Supraleiter,” Physik in Unserer Zeit, 52(2), 80-87, 2021. https://doi.org/10.1002/piuz.202001588

[3] M.J. Lea, P. Frayne, & Y. Mukharsky. “Could we quantum compute with electrons on helium?,” Fortschritte Der Physik, 48(9-11), 1109-1124, 2000. https://doi.org/10.1002/1521-3978(200009)48:9/113.0.co;2-i

[4] S. Kieferle, and S. Markic, “Enabling active participation and inquiry-based learning – inclusive learning settings for a student laboratory for lower secondary school,” Chemkon, 31(1), 21-27, 2023. https://doi.org/10.1002/ckon.202300015

[5] A. Baur, “Fehler, Fehlkonzepte und spezifische Vorgehensweisen von Schülerinnen und Schülern beim Experimentieren,“ ZfDN 24, 115–129, 2018. https://doi.org/10.1007/s40573-018-0078-7

[6] J. Weber, "Mathematische Modellbildung und Videoanalyse zum Lernen der newtonschen Dynamik im Vergleich,", 2022. https://doi.org/10.30819/5566

[7] A. Domínguez, et al. “Gamifying learning experiences: Practical implications and outcomes,” Comput. Educ., 63, 380-392, 2013. https://doi.org/10.1016/J.COMPEDU.2012.12.020

[8] M. Boekaerts, “Engagement as an inherent aspect of the learning process,” Learning and Instruction, 43, 76-83, 2016. https://doi.org/10.1016/J.LEARNINSTRUC.2016.02.001

[9] P. Redmond, et al. “An Online Engagement Framework for Higher Education”, 2018. https://doi.org/10.24059/OLJ.V22I1.1175

[10] J. Fernández-Río, M. Zumajo-Flores, & G. F. Aguilar, “Motivation, basic psychological needs and intention to be physically active after a gamified intervention programme,” European Physical Education Review, 28(2), 432-445, 2021. https://doi.org/10.1177/1356336x211052883

[11] S. S. Feger, S. Dallmeier-Tiessen, P.W. Woźniak, & A. Schmidt, “Gamification in science,” Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 2019. https://doi.org/10.1145/3290605.3300690

[12] F. Forndran, and C.R. Zacharias, “Gamified experimental physics classes: a promising active learning methodology for higher education,” European Journal of Physics, 40(4), 045702, 2019. https://doi.org/10.1088/1361-6404/ab215e

[13] K. Robson, K. Plangger, J. Kietzmann, I.P. McCarthy, & L. Pitt, “Is it all a game? understanding the principles of gamification. Business Horizons,” 58(4), 411-420, 2015. https://doi.org/10.1016/j.bushor.2015.03.006

[14] C. Dichev, and D. Dicheva, “Gamifying education: what is known, what is believed and what remains uncertain: a critical review,” International Journal of Educational Technology in Higher Education, 14(1), 2017. https://doi.org/10.1186/s41239-017-0042-5

[15] N. Zaric, R. Roepke, V. Lukarov, & U. Schroeder, “Gamified learning theory: the moderating role of learners' learning tendencies,” International Journal of Serious Games, 8(3), 71-91, 2021. https://doi.org/10.17083/ijsg.v8i3.438

[16] S. Qiao, S.K.W. Chu, X. Shen, & S.S. Yeung, “The impact of an online gamified approach embedded with self-regulated learning support on students' reading performance and intrinsic motivation: a randomized controlled trial,” Journal of Computer Assisted Learning, 38(5), 1379-1393, 2022. https://doi.org/10.1111/jcal.12684

[17] R. Mitchell, L. Schuster, & J. Drennan, “Understanding how gamification influences behaviour in social marketing,” Australasian Marketing Journal, 25(1), 12-19, 2017. https://doi.org/10.1016/j.ausmj.2016.12.001

[18] C. Wang, J. He, Z. Jin, S. Pan, M. Lafkihi, & X.T. Kong, “The impact of gamification on teaching and learning physical internet: a quasi-experimental study,” Industrial Management & Data Systems, 122(6), 1499-1521, 2022. https://doi.org/10.1108/imds-09-2021-0600

[19] J.-P. Doignon, and J.-C. Falmagne, “Spaces for the assessment of knowledge,” International Journal of Man-Machine Studies 23, 175-196, 1985. https://doi.org/10.1016/S0020-7373(85)80031-6

[20] J.-P. Doignon, and J.-C. Falmagne, “Knowledge Spaces,” Springer-Verlag, Berlin, Heidelberg, and New York, 1999. https://doi.org/10.1007/978-3-642-58625-5

[21] J.-C. Falmagne, “Thirty years of knowledge space theory: the beginning, the core ideas, and the assessment spaces,”. Paper presented at the EMPG Meeting 2015 at Padua, Italy, 01-03 September, 2015.

[22] J. Heller, L. Stefanutti, P. Anselmi, & E. Robusto. “Erratum to: On the link between Cognitive Diagnostic Models and Knowledge Space Theory,” Psychometrika, 81, 250-251, 2016. https://doi.org/10.1007/s11336-015-9494-5

[23] J. Confrey, A.P. Maloney, & A.K. Corley, “Learning trajectories: a framework for connecting standards with curriculum,” ZDM Mathematics Education, 46, 719–733, 2014. https://doi.org/10.1007/s11858-014-0598-7

[24] J. Confrey, R.S. Jones, & G. Gianopulos, “Challenges in Modeling and Measuring Learning Trajectories,” Measurement: Interdisciplinary Research and Perspectives, 13(2), 100-105, 2015. https://doi.org/10.1080/15366367.2015.1055131

[25] J.P. Doignon, and J.C. Falmagne, “Learning Spaces,” Interdisciplinary Applied Mathematics. Springer-Verlag, Berlin, Heidelberg, and New York, 2015.

[26] M. Kambouri, M. Koppen, M. Villano, & J.C. Falmagne, "Knowledge assessment: tapping human expertise by the QUERY routine,” International Journal of Human-Computer Studies, 40, 119-151, 1994. https://doi.org/10.1006/ijhc.1994.1006

[27] M. Schrepp, T. Held, & D. Albert, “Component–based Construction of Surmise Relations for Chess Problems,” In D. Albert & J. Lukas (Eds.), Knowledge Spaces: Theories, Empirical Research, and Applications (pp. 41–66). Mahwah, NJ: Lawrence Erlbaum Associates, 1999.

[28] M. Schrepp, “A Method for Comparing Knowledge Structures Concerning Their Adequacy,” Journal of Mathematical Psychology, 45, 480-496, 2001. https://doi.org/10.1006/jmps.2000.1329

[29] D. Albert, M.D. Kickmeier-Rust, & F. Matsuda, F. «A formal framework for modelling the developmental course of competence and performance in the distance, speed, and time domain,” Developmental Review, 28, 401-420, 2008. https://doi.org/10.1016/j.dr.2008.05.001

[30] R. S. Alsawaier, “The effect of gamification on motivation and engagement,” The International Journal of Information and Learning Technology, 35(1), 56-79, 2018. https://doi.org/10.1108/ijilt-02-2017-0009

[31] M. Nurtanto, N. Kholifah, E. Ahdhianto, A. Samsudin, & F.D. Isnantyo, “A review of gamification impact on student behavioural and learning outcomes,” International Journal of Interactive Mobile Technologies (iJIM), 15(21), 22, 2021. https://doi.org/10.3991/ijim.v15i21.24381

[32] F.L. Khaleel, T.S.M.T. Wook, & A. Ismail, A. “Gamification elements for learning applications,” International Journal on Advanced Science, Engineering and Information Technology, 6(6), 868, 2016. https://doi.org/10.18517/ijaseit.6.6.1379

[33] M.F. Areed, M. A. Amasha, R. A. Abougalala, S. Alkhalaf, & D. Khairy, “Developing gamification e-quizzes based on an android app: the impact of asynchronous form,” Education and Information Technologies, 26(4), 4857-4878, 2021. https://doi.org/10.1007/s10639-021-10469-4

[34] L. Putz, and H. Treiblmaier, “Findings of an experiment: knowledge retention in gamified and non-gamified workshops,” Proceedings of the Annual Hawaii International Conference on System Sciences, 2019. https://doi.org/10.24251/hicss.2019.177

[35] S. Schöbel, A. Janson, & M. Söllner, «Capturing the complexity of gamification elements: a holistic approach for analysing existing and deriving novel gamification designs,” European Journal of Information Systems, 29(6), 641-668, 2020. https://doi.org/10.1080/0960085x.2020.1796531

[36] T.A. Hussein, M. Nat, H.F. Hasan, & A.A. Mahdi, “Development and evaluation of an online gamified assessment environment,” 4th International Conference on Communication Engineering and Computer Science (CIC-COCOS’2022), 2022. https://doi.org/10.24086/cocos2022/paper.744

[37] G. Barata, S. Gama,, J. Jorge, & D. Gonçalves, “Improving participation and learning with gamification,” Proceedings of the First International Conference on Gameful Design, Research, and Applications, 2013. https://doi.org/10.1145/2583008.2583010
[38] I. Hamza, and S. Tóvölgyi, “The effect of gamified e-learning on employees' engagement,” Periodica Polytechnica Social and Management Sciences, 31(1), 80-89, 2022. https://doi.org/10.3311/ppso.19616