International Journal of Science and Mathematics Education
https://doi.org/10.1007/s10763-023-10440-4

Is Immersion in 3D Virtual Games Associated with Mathematical Ability Improvement in Game-Based Learning?

Athanasios Christopoulos[1] · Stylianos Mystakidis[2,3] · Justyna Kurczaba[1] · Mikko-Jussi Laakso[1,4] · Chrysostomos Stylios[5,6]

Received: 20 July 2023 / Accepted: 20 December 2023
© The Author(s) 2024

Abstract

Previous studies have found positive effects of Game-Based Learning for mathematics. While most studies assume that this effect is explained by the presence of flow/immersion during games, this has not yet been established. The aim of the current study is to verify if immersion indeed is associated with mathematical skills improvement when using a Game-Based Learning intervention. This was tested among 59 Greek high school students, using authentic design. After having received a traditional education module, the students were tested and then engaged for four weeks in a desktop-based 3D Virtual Learning Environment where they could play mathematic minigames. They were subsequently re-tested to verify if they showed a significant increase in mathematical skills. The students showed an improvement in their mathematical skills (Cohen's $d = 1.26$), with significant results for functions, geometry, and thinking skills and methods. On the individual level, about half of the students showed a 10% increase in one of the domains (numbers & calculations, functions, geometry, thinking skills and methods, and algorithms and number theory). Immersion was found to be reflected by engagement and presence, but neither one of these aspects was associated with mathematical achievement after the intervention. It is concluded that Game-Based Learning is an effective approach to increasing mathematical skills, yet the underlying mechanisms are not yet understood. The authors discuss several alternative mechanisms based on the literature that can be verified in future studies.

Keywords Engagement · Immersion · Game-based learning · Gamification · Mathematics education · Virtual reality

Extended author information available on the last page of the article

Published online: 04 January 2024

Introduction

Mathematical skills are considered to be important for societies as they contribute to their economic success, and scientific and technological development (Yadav, 2019). For individuals, it is needed to succeed in many other subjects and in a variety of tasks and jobs later in life (Yadav, 2019). However, mathematics is a subject that many adolescents dislike, which is related to students finding it difficult and feeling insecure about their skills (Harun et al., 2021). In fact, about 30% of adolescent students worldwide suffer from mathematics anxiety, characterized by intense feelings of fear and worry during mathematics situations (Klee & Miller, 2019; Organisation for Economic Co-operation and Development, 2013). Girls, in particular, often experience mathematical anxiety, as noted by Klee and Miller (2019). As a consequence, many students are reluctant to involve themselves in mathematics-related study and career paths, as well as Science, Technology, and Engineering (together referred to as 'STEM') that rely heavily on mathematical skills (Li & Schoenfeld, 2019). Moreover, students can show underachievement because of the negative thoughts and emotions associated with mathematics (Fong et al., 2021).

These findings are often explained using the expectancy-value theory. According to this theory, student academic behavior (including choices and persistence) is affected by expectancy beliefs (how well they think they can perform) and by value (how important they think it is; how much they want to do it) (Eccles, 2013; Eccles et al., 1983, p. 83). Several studies have confirmed this theory by showing that students who experience more self-efficacy in mathematics and experience it as more valuable, show more effort and better results (Fong et al., 2021; Michaelides et al., 2019, p. 9; Xu, 2020). Using longitudinal data, Collie et al. (2018) showed that during adolescence, many students over time "switch off", showing increased disengagement from mathematics, associated with decreased expectancies (experiencing the topics as more complicated) and value (mathematics becoming more abstract and less enjoyable). Efforts to increase the perceived value have focused on making the teaching of mathematics more meaningful for students, but these have not been very successful (Harun et al., 2021; Li & Schoenfeld, 2019).

Interestingly, recent studies have focused on the benefits of "Game-Based Learning" (GBL) of mathematics (Fokides, 2018; Wang et al., 2020). The advantage is that students often feel naturally engaged when playing games (Naik, 2015). Previous studies have shown that indeed, positive effects can be obtained when using digital GBL for mathematics or other STEM subjects (Admiraal et al., 2014; Fokides, 2018; Khaleel et al., 2016; Khan et al., 2017). With respect to academic achievement, GBL results in equal or better results compared to traditional learning methods (Fokides, 2018; Kalogiannakis et al., 2021; Khaleel et al., 2016; Lei et al., 2022; Tsai & Tsai, 2020). With respect to engagement and motivation, GBL usually outperforms traditional methods, especially for girls (Admiraal et al., 2014; Kalogiannakis et al., 2021; Khan et al., 2017).

However, a recent literature review of Krath et al. (2021) revealed that, although these outcomes are well-established, there is a research gap with respect to finding evidence for the assumed mechanisms. While studies are carried out from a range of theoretical perspectives, the 'flow' theory is one of the most dominant ones (Krath et al., 2021). The experience of 'flow' refers to a state during which people work with full capacity, intrinsically motivated and focused (Nakamura & Csikszentmihalyi, 2014, p. 240). This concept was introduced by Csikszentmihalyi who first recognized it in artists who could work on a painting with full concentration, disregarding feelings such as hunger and fatigue (Csikszentmihalyi & Csikszentmihalyi, 1988, p. 29; Nakamura & Csikszentmihalyi, 2014, p. 239). Research has demonstrated that people can experience flow in a variety of activities (including artistic, sports, and games), usually when they are challenging and have proximate goals with immediate feedback (Nakamura & Csikszentmihalyi, 2014, p. 240). It is assumed that adding play in this way may compensate for a lack of positive expectations and/or perceived value regarding mathematics. Indeed, it has been found that with GBL students can experience flow and their possible maths anxiety fades into the background (Wang et al., 2020). Overall, the structured steps, short-term goals, and direct feedback that are relevant from a flow-perspective, seem to also be the most important elements identified when studied from other theoretical frameworks (e.g. self-determination theory) that are thought to increase motivation and engagement (Krath et al., 2021).

A separate line of research has focused on the experience of flow in digital games, especially in 3D games that provide players with a rich experience (Boyle et al., 2012; Hookham & Nesbitt, 2019; Marto & Gonçalves, 2022). Within these studies, many different labels have been given to flow, with some different definitions, including 'engagement', 'immersion', and 'presence' (Agrewal et al., 2020; Christopoulos et al., 2018; Michailidis et al., 2018; Mount et al., 2009). The most important distinction is that players can feel engaged in a game in a flow-like experience (i.e. being focused and intrinsically motivated) while also aware that they are playing a game; yet they can also feel present in the virtual environment (Michailidis et al., 2018; Mount et al., 2009). Mount et al. (2009) defined the former (engagement) as "the acceptance of presented narratives and the suspension of disbelief, and strategies, actions and technologies surrounding it" and the latter (presence) as "an individual's experience of being on one place or environment, whilst being physically situated in another" (p. 53), while seeing both concepts as part of the larger concept 'immersion'. As such, immersion is considered to be a synonym for flow (Michailidis et al., 2018). These definitions will be followed in the current study.

This conceptual uncertainty extends even to the definition of virtual environments per se (Girvan, 2018). Metaverse platforms, another terminology in the field of virtual worlds, are seen to provide potential benefits for learning (Hwang et al., 2023). However, despite the apparent potential of these technologies, empirical research is lacking on their application in education (Asiksoy, 2023). To illustrate this need for empirical research, we can look towards the field of GBL. This approach has gained considerable attention recently, but the existing studies often overlook key elements such as immersion. For instance, the systematic review of Laine and Lindberg (2020) discusses how to design engaging educational games based on previous studies, but the review of Checa and Bustillo (2020) indicates that these studies often fail to measure the effect of immersion in the learning process.

In this context, the study by Özhan and Kocadere (2019) stands out, showing that in a mixed learning environment where GBL is combined with traditional methods, immersion can partly explain academic success in undergraduate university students. While not focused on mathematics, nor on adolescent students, the study of Özhan and Kocadere (2019) is interesting for two reasons. First, it used an authentic study design as educational institutes tend to primarily use traditional methods such as lectures and individual or group exercises, and are more likely to combine these with GBL (i.e. choosing a mixed or "blended" learning environment) rather than opting for complete substitution (Jing et al., 2015) and, second, it provides a first confirmation for the assumption that immersion is a factor through which GBL contributes to academic performance.

Therefore, the aim of the current study is to verify whether immersion explains the effects of GBL on mathematics achievement among high school students. Following the study Özhan and Kocadere (2019) this was also done using a naturalist study design, yet slightly different. In order to be able to measure the effects of GBL, students first received traditional mathematics classes on various topics, were then tested, and thereafter played 3D mini games. This allowed us to more specifically address the added value of GBL and whether this was mediated by immersion. In addition, given the fact that more girls than boys experience mathematical anxiety (Klee & Miller, 2019), we controlled for gender differences. Finally, it is also known that student's use of learning strategies affects their mathematics achievement (Altun & Erden, 2013; Sun et al., 2018). For this reason, learning strategies were also taken into account. As such, our research hypotheses are:

H1. After the GBL intervention, students will on average show further improvements in their mathematical skills.

H2. The improvement in mathematical skills will be predicted by immersion.

Materials and Methods

Participants and Setting

The study was carried out in collaboration with a private tutoring school in Greece. To this end, it should be noted that: *(a)* the participants were concurrently attending their regular school and *(b)* they came from various local schools, each with its unique academic environment, which could potentially contribute to the diversity in their learning experiences and outcomes. The integration of this study into their broader educational context aimed to supplement their regular school curriculum, particularly in mathematics, without interfering with their usual school schedule.

The parents/guardians of the students who expressed an interest in participating were asked to countersign an informed consent form which included information about the following: *(a)* warnings of known possible side-effects that might occur while using the VR environment, *(b)* information about the collection and use of students' data and *(c)* their right to withdraw from the study at any point without any consequence. In total, 59 (33 boys, 26 girls) high school students (14 and 15 years old) participated in the study.

Instructional Design

Students were asked to engage with the desktop-based 3D Virtual Learning Environment outside the classroom context (i.e. on voluntary basis). The mathematics minigames that were available were designed and developed by the research team based on their appropriateness for age and mathematical skills, and in line with what is known from the literature to help increase immersion (Laine & Lindberg, 2020). The 14 minigames are described in Table 1, following the Learning Mechanics–Game Mechanics model developed by Arnab et al. (2014) to theoretically describe the learning and game mechanics of educational games.

Research Design

Procedure

At baseline, students filled out their demographics and computer use. They also were given an assessment form (pre-test) to assess their math skill. They further completed the survey to assess their learning strategies. Thereafter followed a four-week period during which the students played the 3D minigames for at least 45 min a week. The system was registering whether students actually played the games and how many times. After playing each game for the first time, the students rated their enjoyment on a 5-point Likert scale. After this period, the post-assessment took place during which the students completed a survey about immersion and again were tested on their mathematical skills. This second test was similar to the first test, using parallel, but different problems to avoid a test–retest effect. The experiment concluded with an anecdotal focus group discussion where teachers and students could discuss with the principal researcher the benefits and limitations of this alternative instructional approach. All sensitive data were handled in accordance with the General Data Protection Regulation (GDPR).

Instruments

Learning strategies were measured using five subscales of the Learning Strategies scale of Berger and Karabenick (2011): *(1)* rehearsal, *(2)* planning, *(3)* monitoring, *(4)* regulation and *(5)* help seeking. In total, these were measured using 21 items on a 5-point Likert scale from 1 = *Not at all true of me* to 5 = *Very true of me*. However, because Berger and Karabenick found that not all items showed (sufficient) loadings on the intended factors, a Confirmatory Factor Analysis (CFA) was carried out using JASP 14.1. Based on the poor fit of the initial model ($\chi^2(179) = 361.97$, CFI = 0.695, RMSEA = 0.132), it was decided to discard 6 items with factor loadings < 0.40. This improved the model fit ($\chi^2(80) = 112.51$, CFI = 0.924, RMSEA = 0.083). The items and their factor loadings are displayed in Table 2. The subscales were subsequently

Table 1 Key elements of the Learning-Game Mechanics (Arnab et al., 2014) of the 3D minigames ($n = 14$)

Minigames	Learning Mechanics		Game Mechanics	
	Abstract	Concrete	Abstract	Concrete
Multiplication Board [Singleplayer]	Guidance	Repetition Assessment Feedback	Realism Feedback	Quick feedback Infinite gameplay Assessment
Fractions Line Up [Singleplayer]	Guidance	Analyse Experimentation Assessment Feedback	Fun Challenge Realism Feedback	Quick feedback Time pressure Assessment
Declare A Fraction War [Singleplayer]	Competition	Analyse Assessment Feedback	Fun Challenge Realism Feedback	Time pressure Assessment Quick feedback
Declare A Fraction War [Multiplayer]	Competition	Analyse Assessment Feedback	Fun Challenge Realism Feedback	Time pressure Game turns Assessment Quick feedback
Fraction Hopscotch [Singleplayer]	N/A	Analyse Assessment Feedback	Fun Challenge Pavlovian interactions Realism Feedback	Movement Assessment Quick feedback
Quadratic Function Analyzer [Singleplayer]	Instructional	Analyse Assessment Feedback	Challenge Realism Feedback	Assessment Quick feedback
Linear Function Analyzer [Singleplayer]	Instructional	Analyse Assessment Feedback	Challenge Realism Feedback	Assessment Quick feedback
Radians Conversion Code Colouring [Singleplayer]	Guidance	Analyse Action / Task Assessment Feedback	Fun Challenge Realism Feedback	Movement Assessment Quick feedback
Degrees Conversion Code Colouring [Singleplayer]	Guidance	Analyse Action / Task Assessment Feedback	Fun Challenge Realism Feedback	Movement Assessment Quick feedback
Polygon Party [Singleplayer]	N/A	Analyse Assessment Feedback	Feedback	Assessment Quick feedback
Cartesian Battleship [Multiplayer]	Plan Competition	Analyse Feedback	Fun Challenge Strategy / Planning Competition Feedback	Game turns Capture / Eliminate Time pressure Quick feedback
The Scale Factor [Singleplayer]	Guidance	Analyse Simulation Assessment Feedback	Fun Challenge Realism Feedback	Assessment Simulate / Response Quick feedback

Table 1 (continued)

Minigames	Learning Mechanics		Game Mechanics	
	Abstract	Concrete	Abstract	Concrete
Fill'er Up—Liquid Measurement [Singleplayer]	Guidance	Analyse Simulation Assessment Feedback	Fun Realism Competition	Simulate / Response Quick feedback
Connecting Europe [Singleplayer]	N/A	Analyse	Challenge Feedback	Quick feedback Assessment

computed by taking the average of the associated items. We found acceptable to good internal consistencies with $\alpha = 0.79$ for Rehearsal, $\alpha = 0.75$ for Planning, $\alpha = 0.90$ for Monitoring, $\alpha = 0.82$ for Regulation and $\alpha = 0.82$ for Help-seeking.

Immersion was measured using the immersion and engagement items of the User Experience in Immersive Virtual Environment scale (Tcha-Tokey et al., 2016). These were 10 items measured on a 5-point Likert scale from 1 = *Strongly Disagree* to 5 = *Strongly Agree*. Because the developers of the scale had not verified the factor structure, an Exploratory Factor Analysis (EFA) was carried out in SPSS version 28.01. Principal axis factoring was used, and Promax rotation to facilitate interpretation. The scree plot suggested a 2-factor structure, that explained 65.3% of the variance in the data. The rotated factor solution is demonstrated in Table 3. The items loading on the first factor were all about involvement with the game in a flow-like manner; the items loading on the second factor were about feeling really in the game. This fitted well with the definitions of 'engagement' and 'presence', as presented in the Introduction, and therefore the factors were labelled accordingly. These two subscales of immersion were computed by taking the average of the associated items. The scales showed good internal consistency, with Cronbach's $\alpha = 0.89$ for engagement and $\alpha = 0.81$ for presence.

Mathematical skill was measured using 2 parallel tests at pre- and post-measurement. Each consisted of 36 problems covering the following topics: *(a)* numbers and calculations, *(b)* functions, *(c)* geometry, *(d)* thinking skills and methods, and *(e)* algorithms and number theory. The test score was calculated by first calculating the percentage correct for each topic, and then by taking the overall average of this for the 5 topics together. This was done by the teachers in charge as part of their routine assessment on students' progress monitoring.

Data Analysis

It was established that the variables showed approximately normal distributions, with the z-values for the skewness and kurtosis below 3.29 (Kim, 2013). Changes in mathematical test scores were compared using paired sample *t*-tests. As recommended to more fully grasp the practical value of changes, we also established on the individual level whether there were changes that would translate to real-world value (e.g. Doyle et al., 2022). As the current study was concerned with test scores

Table 2 Factor structure of the learning strategies

Factor	Item	Estimate	*SE*	*z*-value
Rehearsal	When I study math, I memorize what I need to learn by repeating it over and over to myself	0.85	0.10	8.36**
	I study math by going over the formulas or definitions in order to memorize them	0.69	0.10	7.06**
	I study math by doing the practice problems over and over again to memorize them	0.46	0.11	4.22**
Planning	When I learn new topics in math, I first figure out the best way to study	0.58	0.09	6.29**
	Before I study math, I set goals for myself to help me learn	0.64	0.10	6.37**
	I plan how I am going to study new math topics before I begin	0.49	0.12	4.22**
	Before I study math, I plan how much time I will need to learn a topic	0.45	0.11	4.12**
Monitoring	When I'm studying math, I test myself to see whether I know the material	0.81	0.10	8.39**
	I check whether I have learned what I am studying in math	0.90	0.10	9.09**
Regulation	If I get confused with something I'm studying in math, I go back and try to figure it out	0.86	0.11	7.98**
	If the math I am studying is difficult to learn, I slow down and take my time	0.69	0.10	6.92**
	If I'm having trouble solving math problems I try other ways to solve them	0.53	0.09	5.70**
Help seeking	If I don't understand something in math I ask my teacher for help	0.71	0.10	7.12**
	If I don't understand something in math I ask other students for help	0.68	0.11	6.32**
	If I don't understand something in math I ask for help to better understand general ideas or principles	0.61	0.10	6.03**

**$p < 0.001$

Table 3 Factor Structure of the immersion items

Item	Factor	
	Engagement	Presence
I felt stimulated by the virtual environment	0.73	-0.10
I become so involved in the virtual environment that I was not aware of things happening around me	-0.28	0.65
I identified to the character I played in the virtual environment	0.59	-0.03
I become so involved in the virtual environment that it is if I was inside the game rather than manipulating a gamepad and watching a screen	-0.01	0.74
I felt physically fit in the virtual environment	0.18	0.87
I got scared by something happening in the virtual environment	0.02	0.63
I become so involved in the virtual environment that I lose all track of time	0.80	-0.10
The visual aspects of the virtual environment involved me	0.79	0.07
The sense of moving around inside the virtual environment was compelling	0.77	0.09
I was involved in the virtual environment experience	0.85	0.03

that could maximumly reach 100 (percentage correct), we considered a 10-point difference in a mathematical domain to reflect a 'real' or "practically significant" change. Associations between the predictors and test scores were determined by Pearson correlations. We subsequently carried out two regression analyses. One used the test scores after GBL as the dependent variable and the other one used the difference scores between the two tests as the dependent variable. Three models were used, to verify if adding the immersion variables after controlling for gender and learning strategies would be significant. Moreover, we tested a third model for sake of completeness to more fully take into account possible gender effects, by adding the interaction between gender and immersion (i.e. moderation). Gender was dummy coded with 0 for boys and 1 for girls.

Results

Descriptive Statistics

Prior to the conduct of the study, participants were asked about their normal computer use. All students regularly used a computer. Precisely, 96.6% use computers for learning 1 to 10 h per week, and the remaining 3.4% use it even more. Most students (86.4%) also use computers for gaming. The students were 14 or 15 years old (59.3% and 40.7%, respectively). The average frequency of playing the educational math games ranged from 3.71 to 6.57 h, with a mean of 5.12 h ($SD = 0.62$). Still, the average rating of the games ranged from 2.64 to 3.71, with a mean of 3.40 ($SD = 0.23$), indicating that even though the students choose to play the games repeatedly, they tended to be neutral about the likability of the games. Looking at the different games separately, the most popular games were the multi-player "Declare a Fraction War" (average rating = 3.68) and the single-player "Fraction Hopscotch", that both were played on average almost 6 times. The average rating of each game was above 3 (the scale midpoint). Ratings and gaming frequency were not associated with the final test results ($r = 0.03$, $p = 0.839$ and $r = 0.004$, $p = 0.978$ respectively), neither with the improvement compared to initial testing ($r = -0.11$, $p = 0.407$ and $r = -0.13$, $p = 0.310$ respectively).

Mathematical Skill

Table 4 shows the descriptive statistics of the test scores at baseline and after having played the games for 4 weeks, as well as the paired sample *t*-tests. There was a significant improvement in mathematical skill, as indicated by the significant increase in the overall test scores. The improvements in functions, geometry, and thinking skills and methods were significant, but the test scores for numbers and calculations as well as algorithms and number theory did not show a significant improvement. The increase in the overall test scores was not just significant ($t(58) = 9.65$, $p < 0.001$), but also showed a large effect size, with Cohen's $d = 1.26$. Also, it is worth noting that, with respect to real world value, about half of the sample (45.8%)

Table 4 Descriptive statistics and paired sample *t*-tests for the mathematical test results

	Baseline	After GBL	*t*-test	
Test	*M* (*SD*)	*M* (*SD*)	*t*(58)	*p*
Numbers and calculations	52.98 (10.92)	53.68 (11.72)	1.02	0.313
Functions	44.20 (12.27)	49.59 (11.10)	7.70	< 0.001
Geometry	45.86 (10.66)	50.51 (12.23)	6.65	< 0.001
Thinking skills and Methods	51.29 (12.87)	53.68 (12.79)	3.42	0.001
Algorithms and Number theory	49.61 (10.75)	50.36 (10.68)	1.63	0.109
Total	48.79 (7.18)	51.56 (7.42)	9.65	< 0.001

showed a 10% increase in one of the mathematical domains. As such, the first hypothesis can be fully confirmed: after the GBL intervention, the students showed further improvements in their mathematical skills.

Learning Strategies and Immersion

Table 5 shows the descriptive statistics for the learning strategies, game immersion and their correlations with the test results at baseline and after the GBL. In contrast to the expectations, almost no significant associations with the test results were found. The only effects that were significant indicated that students who reported more Rehearsal and Planning, had received lower grades. We further tested if there were gender effects on the test results, but girls and boys had about equal test scores, both at baseline ($M_{girls}=48.49$, $SD=7.64$ and $M_{boys}=49.02$, $SD=6.92$, $t(57)=-0.28$, $p=0.780$) and after the GBL intervention ($M_{girls}=51.24$, $SD=7.68$ and $M_{boys}=51.82$, $SD=7.33$, $t(57)=-0.30$, $p=0.769$).

In line with these findings, the regression analysis revealed no effects of immersion on the test results after GBL (Table 6). In the first model, with baseline scores, gender, and learning strategies used as predictors, only mathematics at baseline had

Table 5 Descriptive statistics and correlations with test results for Learning Strategies and game immersion

	Descriptive		*Correlations*	
Test	*M*	*SD*	Maths Test 1 (Baseline)	Maths Test 2 (After GBL)
Rehearsal	3.16	0.73	-0.26*	-0.20
Planning	2.99	0.62	-0.32*	-0.28*
Monitoring	2.58	0.91	0.23	0.24
Regulation	2.75	0.76	0.16	0.21
Help seeking	3.05	0.74	0.04	0.01
Engagement	3.37	0.78	0.15	0.11
Presence	2.86	0.63	0.06	0.09

* $p < 0.05$

Table 6 Mathematics test scores after GBL regressed on gender, learning strategies and immersion

Predictor	Model 1			Model 2			Model 3		
	B	*SE*	β	*B*	*SE*	β	*B*	*SE*	β
Mathematics at baseline	1.00	0.05	0.97**	1.01	0.05	0.98**	1.01	0.05	0.98**
Gender	0.08	0.61	0.01	0.19	0.66	0.01	-0.37	4.15	-0.02
Rehearsal	0.66	0.47	0.06	0.65	0.48	0.06	0.64	0.49	0.06
Planning	-0.01	0.58	0.00	0.15	0.61	0.01	0.13	0.63	0.01
Monitoring	-0.67	0.63	-0.08	-0.83	0.66	-0.10	-0.84	0.67	-0.10
Regulation	1.27	0.70	0.13	1.27	0.72	0.13	1.30	0.75	0.13
Help seeking	-0.27	0.43	-0.03	-0.31	0.44	-0.03	-0.31	0.46	-0.03
Engagement				-0.18	0.43	-0.02	-0.26	0.52	-0.03
Presence				0.56	0.52	0.05	0.62	0.74	0.05
Gender*engagement							0.25	0.89	0.06
Gender*presence							-0.11	1.05	-0.02

a significant effect ($\beta = 0.97$, $p < .001$), $R^2 = .92$, $F(7, 51) = 84.66$, $p < .001$. Adding presence and engagement as predictors in Model 2, revealed that this did not significantly improve the prediction of the test results, $\Delta R^2 < .01$, $F(2, 49) = 0.62$, $p = .543$. Testing for potential moderation of the immersion effects by gender in Model 3 also did not result in significant model improvement, $\Delta R^2 < .01$, $F(2, 47) = 0.04$, $p = .960$. We then tested the same models of predictors for the test improvement scores. This analysis revealed no significant results (therefore, these are not presented for reasons of parsimony). It is noted that improvement was also not associated with initial test scores ($\beta = 0.01$, $p = .954$).

We also tested whether students who had shown a change of 10% of improvement in a mathematical skill differed from those who did not show such an improvement in the presence and engagement they reported. This was not the case; students with a domain improvement of 10% reported an average engagement of 3.24 ($SD = 0.78$) and presence of 2.81 ($SD = 0.42$) and this was similar for students who had not achieved a 10% increase in a mathematical domain with an average engagement of 3.48 ($SD = 0.77$) and presence of 2.90 ($SD = 0.77$), $t(57) = -1.21$, $p = .232$ and $t(57) = -0.51$, $p = .616$ respectively. It is concluded that the second hypothesis should be rejected: immersion does not predict improvement in test scores.

Discussion

The aim of the current study was to examine whether immersion could explain the effect of GBL on mathematical skills. For this purpose, students were asked to involve themselves in playing 3D minigames that covered mathematical topics. Although their mathematical test results subsequently showed the expected improvement, this was not explained by immersion. We also found no gender effects and little effects of learning strategies.

The fact that we found no evidence for immersion explaining increases in mathematical skill, is an important finding. After all, most studies interested in GBL assume that immersion/flow is the mechanism that explains the effects of GBL (Krath et al., 2021). The current study results seem to be in contrast with the results of Özhan and Kocadere (2019) who were able to demonstrate an association between flow and academic outcomes. However, there are some important differences between their study and ours. First of all, their study was based on research among undergraduate students taking a university course. As such, both the age and topic were different. More importantly, in their study, all academic learning possibilities were presented simultaneously. As such, it was not possible to verify to what extent it was indeed immersion with the games that explained academic results, or topic-motivation and general engagement. In fact, they did not find a significant effect of what they labelled "emotional engagement" with GBL (i.e. feeling interested, liking it, not being bored), and the effect of immersion on academic achievement was not significant once general (topic) motivation was controlled for. Even though the authors presented this result as the effect of flow being mediated by motivation, their study did not allow for the testing of the direction of causality. As such, it might be the case that topic-motivation simply positively affected game-immersion.

In the current study, we were better able to disentangle the effects of the traditional learning methods and GBL by introducing GBL after the effects of the traditional method (potentially affected by motivation for mathematics) had taken place. As such, we were able to measure the possible effects of game immersion. Moreover, our conclusion that immersion was not the mechanism through which GBL helps students improve mathematical skills is strengthened by the finding that the game-based ratings were also not associated with improvements in mathematical skills. After all, from flow theory, immersion is closely related to motivation, as immersion/flow is seen as an intrinsic process of persistent motivation (Nakamura & Csikszentmihalyi, 2014, p. 216). Interestingly, the most popular games were in the domain of Numbers & Calculations, and this domain did not show a significant improvement.

This means that there must be a different explanation of why GBL helps improve mathematical skills in students. The seminal work of Nuthall (1999, 2004, 2005, 2007) provides us with several useful clues. Nuthall's granular analysis of individual learning processes revealed that learning occurred as a sequence of events that each build on the previous one (Nuthall, 1999, 2005). Learning was not determined by the specific learning activities; instead, it depended on whether the information made sense to the learner at that moment determined learning. This is in line with what many educators know as Vygotsky's (1978, p. 86) "zone of proximal development". Nuthall found that students learned a concept when they had been exposed to the complete set of information three times (Nuthall, 1999, 2004, 2005, 2007). Educational games are usually developed in a way that the student is indeed confronted with basic information, which is repeated as the player continues, only increasing the level if the player is successful (Laine & Lindberg, 2020; Moyer-Packenham et al., 2019). This is a big advantage compared to whole-classroom lessons that are more difficult to individualize. GBL circumvents this limitation by adapting to the learner's progress, providing immediate feedback, and offering varying levels of difficulty to suit individual learning trajectories. This approach not only fosters a deeper understanding of concepts but also promotes engagement and motivation, as students are more likely to be invested in learning activities that resonate with their personal learning journey. Besides, the iterative nature of games, where concepts are revisited and reinforced through repeated exposure, aligns seamlessly with the cognitive principles of learning.

A second finding of Nuthall (2005, 2007) was that those students who showed the most learning created more learning experiences by asking questions, talking to others about the concepts, and so on. The development of these behaviours is influenced both by the student's background, such as the presence of similar behaviours in their home environment, and by their personal expectations. Students who view themselves as incapable often inadvertently reduce their own chances of engaging in additional learning opportunities (Nuthall, 2005). Adding GBL to the normal educational program, might create learning opportunities that some students will not self-create otherwise. Even though motivation and frequency of playing the games were not associated with the test results, all students in our study played each game at least once, and we found that on average each game was played at least three times. This means that the students had more learning opportunities.

A third mechanism that we wish to suggest, is that the GBL may significantly increase mathematical fluency. In our study, we opted for percentage correct as the metric for mathematics skill scores. This measure, an intuitive and widely accepted gauge of test performance, allows for the assessment of individual progress without the influence of comparative scores. This approach is well-established; as Binder remarks: "We all grew up in a percentage correct world" (Binder, 2003, p. 14). However, this method has its own limitations. It does not comprehensively capture all dimensions of mathematical skills, such as fact knowledge, speed, strategy, and solution flexibility, which are integral components of mathematical fluency (Akkan, 2021; Cartwright, 2022). Furthermore, while the test in our study was not designed to be time-constrained, the 45-min limit and potential social pressure from peers completing the test could have influenced student performance. Mathematical fluency is linked to a flexible and conceptual approach to mathematics, fostered by encouraging students to experiment with numbers and mathematical concepts (Boaler, 2019; Cartwright, 2022). Digital games, by their nature, promote quick thinking, creativity, and strategic planning, and provide immediate feedback. These elements are conducive to developing the mathematical mindset that is essential for enhanced mathematical fluency (Hulse et al., 2019).

The fact that we found no gender differences in our study is promising. Although mathematical anxiety seems to be more prevalent in girls and females tend to be underrepresented in STEM educational and career tracks, do they not necessarily have less potential or lower mathematical/STEM abilities (Ghasemi & Burley, 2019). In other words, girls have just as much potential, and mathematics education should, therefore, be appropriate for each gender so that this potential can be utilised and not lost due to a lack of self-efficacy or interest. The fact that the effect of GBL was not affected by gender means that it can be used with equal success for boys and girls (it is noted that in the current study, there were no students who identified themselves as non-binary).

With respect to the lack of findings on learning strategies, this is somewhat puzzling. Even though it was only included as a control variable, one would expect that more correlations would be found between the learning strategies and especially the initial test results. To this end, our measurement of learning strategies was limited because we only used students' self-reported strategies on closed-questions after having taken the baseline test. Other researchers have suggested using an open-question format as students then will report about strategies that were most salient for them (Dirkx et al., 2019; Karpicke et al., 2009). The fact that we found negative correlations between rehearsal and planning and the test results might be because students with more mathematical difficulty felt more need to practice. With respect to monitoring, regulation, and help-seeking, it might be that students were unable to correctly reflect on how frequently they had actually used these strategies. It also seemed to be that in this particular school, most of the student learning was organized within the classroom. The teachers seemed to be involved and focused on providing proactive help to students in need. As such, self-regulated learning might have had less impact compared to in other situations.

Implications

The study has some important implications. First of all, the finding that mathematical skills increased by GBL, is in line with previous research findings (Admiraal et al., 2014; Fokides, 2018; Khaleel et al., 2016; Khan et al., 2017) and forms a further confirmation of the positive effects of GBL on mathematical skills. Importantly, while usually only measured on the average level, we also investigated if changes were achieved that would have real-world value for individual students. Although somewhat arbitrary, a 10-point difference on a 100-point scale is in line with what other researchers have identified as noteworthy (Cooper, 2010; Field et al., 2019) and translates to a 1-point grade difference if reduced to a 10-point scale that is also frequently used in many countries (Jakaitienė et al., 2021; Tran & Nguyen, 2021; UK National Recognition Information Centre, 2004). Given the large effect that was obtained using only a 4-week period on a completely voluntary basis makes GBL a great addition to the traditional learning program that seems to be more efficient compared to tutoring, as this usually shows smaller effects and can be a more time-consuming and expensive option (Alegre et al., 2019).

Another implication of our study result relates to the theory about the mechanisms through which GBL works. While we acknowledge the effect that immersion can have on learning outcomes, it was not identified as the working mechanism for GBL. The alternative explanations above mentioned, and maybe other ones, need to be addressed in future studies. Given the evidence for the effects of GBL on mathematics, identifying the mechanisms seems to be an important next step. This is particularly true to find ways to further improve GBL, as will be explained in the next section.

Although it was not a primary aim of the current study, we also wish to stress the theoretical implications of our findings with respect to the dimensions of immersion. As explained in the introduction, 'immersion' and 'flow' are synonyms, with the former label used more frequently in the arts and motivational literature (Csikszentmihalyi & Csikszentmihalyi, 1988 p. 32; Nakamura & Csikszentmihalyi, 2014, p. 79) and the latter more specifically in the educational gaming literature (Agrewal et al., 2020; Michailidis et al., 2018; Mount et al., 2009). Yet, various labels with different definitions are used and this can be confusing for researchers (Agrewal et al., 2020; Michailidis et al., 2018; Mount et al., 2009). In the current study, we found that the theoretical distinction between engagement as immersion by acceptance of and involvement with the presented content and presence as a feeling of being in the game world previously suggested by Mount et al. (2009) was indeed confirmed. As such, future studies can use these dimensions as well and further clarify how these aspects of immersion are relevant for learning processes or related concepts, such as motivation. Even though we could not confirm the effect of engagement and presence on the mathematical improvements, these might for example be factors that affect whether students are willing to continue playing certain games over longer periods of time.

Limitations and Future Work

The limitations of the current study also need to be mentioned as they provide further suggestions for future research.

First of all, the study was limited by the use of only one cohort. Future research could verify if similar results will be found. Most importantly, it is noteworthy that the students were willing to play the games, even if they were not rated extremely high in terms of likeability. This might have been related to the students' having positive relationships with their mathematics teachers. These were three different teachers (two males and one female) that seemed to have different approaches and personalities but came across as quite passionate about mathematics and establishing positive teacher-student relationships. While this impression is highly subjective, it might be addressed more objectively in future studies to what extent teacher-student relationships play a role. These kinds of aspects may affect students' willingness to cooperate and could be addressed in future studies (Semeraro et al., 2020). Indeed, the fact that there was room for improvement with respect to the attractiveness of the games also means that future researchers should try to improve this. As discussed by Collie et al. (2018), students in this age group may become disengaged or lose interest in mathematics. If students find positive experiences through GBL, this might be prevented, giving students a positive experience that increases their self-efficacy. Therefore, if this can be stimulated even further, that could be particularly valuable results (Fong et al., 2021; Michaelides et al., 2019, p. 11; Xu, 2020).

Second comes the lack of a control group. On the one hand, this allowed for a strong design that was authentic. As such, we were able to study the use of GBL in a way that it would also be implemented by schools, offering it to a whole classroom as an addition to the classic teaching approach (with a textbook, lectures, etc.). On the other hand, it did not allow for a comparison with how achievement without being offered GBL or other types of additional learning opportunities. This could be addressed in future studies. For example, this could be done by randomization of an intervention and control group at the school level. Having found initial evidence of positive effects, this means that a further investment for a larger scale study that this requires is worthwhile.

Another limitation that needs to be addressed is the fact that the games can be further optimized in terms of learning content. After all, we failed to find significant improvements in numbers and calculations and algorithm and number theory. Understanding the underlying mechanisms through which GBL works can probably be informative in this respect. For example, if games are helpful in improving mathematics because they tap into the skills necessary for mathematical fluency, the learning mechanisms could be brought even more in line with these. Furthermore, while we did not find the expected effects of learning strategies on achievement, this (as already explained) might be related to our assessment of learning strategies and the limited self-regulated learning that seemed to take place. Previous research has revealed that high school students do not yet seem to use the most effective strategies and a need is identified to help students select and combine, as well as apply strategies to improve their self-studying (Dirkx et al., 2019; Fritz et al., 2007). Future studies could address this issue using GBL with games that are designed to stimulate students in applying certain learning strategies.

A final limitation that needs to be considered for future research is the nuanced representation of mathematical concepts in educational games. While our study primarily focused on the overall impact of these games on student engagement and achievement, further exploration into how mathematical ideas are presented within these games could be beneficial. This exploration could contribute to a deeper understanding of how game design influences learning outcomes in mathematics.

Acknowledgements The authors would like to thank Mr. Apostolos Koutris, Mr. Vaggelis Valomenos, and Ms. Marianna Tsaprouni for supporting the conduct of this study as well as all the students and parents who participated in this initiative.

Funding Open Access funding provided by University of Turku (including Turku University Central Hospital). The work of the first, third, and fourth author has been supported by the Erasmus + project "MATHESIS" (Project code: 2020-1-RO01-KA201-080410).

Data Availability The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics Approval Statement All procedures performed in studies involving human participants were following the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the research as well as from their legal guardians/parents.

Participation in the Study Participants personal information were collected but the right to withdraw at any time was granted.

CRediT **Conceptualization**: A.C., S.M.; **Methodology**: A.C.; **Formal analysis and investigation**: A.C.; Writ**ing—original draft preparation**: A.C.; **Writing—review and editing**: A.C., S.M.; **Funding acquisition**: AC, M.J.L.; **Software**: J.M.K; **Supervision**: M.J.L, C.S.

Conflict of Interest All authors declare that they have no conflicts of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Admiraal, W., Huizenga, J., Heemskerk, I., Kuiper, E., Volman, M., & Ten Dam, G. (2014). Gender-inclusive game-based learning in secondary education. *International Journal of Inclusive Education, 18*(11), 1208–1218. https://doi.org/10.1080/13603116.2014.885592

Agrewal, S., Simon, A. M. D., Bech, S., Bærentsen, K. B., & Forchammer, S. (2020). Defining Immersion:: Literature Review and Implications for Research on Audiovisual Experiences. *Journal of the Audio Engineering Society*, *68*(6), 404–417. https://doi.org/10.17743/jaes.2020.0039

Akkan, Y. (2021). Mathematics teachers' understanding of the concept of mathematical fluency. *Turkish Journal of Mathematics Education, 2*(2), 84–109.

Alegre, F., Moliner, L., Maroto, A., & Lorenzo-Valentin, G. (2019). Peer tutoring and mathematics in secondary education: literature review, effect sizes, moderators, and implications for practice. *Heliyon*, *5*(9), Article e02491. https://doi.org/10.1016/j.heliyon.2019.e02491

Altun, S., & Erden, M. (2013). Self-regulation based learning strategies and self-efficacy perceptions as predictors of male and female students' mathematics achievement. *Procedia-Social and Behavioral Sciences, 106*, 2354–2364. https://doi.org/10.1016/j.sbspro.2013.12.270

Arnab, S., Lim, T., Carvalho, M. B., Bellotti, F., de Freitas, S., Louchart, S., Suttie, N., Berta, R., & De Gloria, A. (2014). Mapping learning and game mechanics for serious games analysis. *British Journal of Educational Technology, 46*(2), 391–411. https://doi.org/10.1111/bjet.12113

Asiksoy, G. (2023). Empirical studies on the metaverse-based education: A systematic review. *International Journal of Engineering Pedagogy, 13*(3), 120–133. https://doi.org/10.3991/ijep.v13i3.36227

Berger, J.-L., & Karabenick, S. A. (2011). Motivation and students' use of learning strategies: Evidence of unidirectional effects in mathematics classrooms. *Learning and Instruction, 21*(3), 416–428. https://doi.org/10.1016/j.learninstruc.2010.06.002

Binder, C. (2003). Doesn't everybody need fluency? *Performance Improvement*, *42*(3), 14–20. http://www.drcarlbinder.org/Need_Fluency.pdf. Accessed 23 Dec 2023

Boaler, J. (2019). Developing mathematical mindsets: The need to interact with numbers flexibly and conceptually. *American Educator*, *42*(4), 28–40. https://files.eric.ed.gov/fulltext/EJ1200568.pdf. Accessed 23 Dec 2023

Boyle, E. A., Connolly, T. M., Hainey, T., & Boyle, J. M. (2012). Engagement in digital entertainment games: A systematic review. *Computers in Human Behavior, 28*(3), 771–780. https://doi.org/10.1016/j.chb.2011.11.020

Cartwright, K. (2022). What does true mathematical fluency look like in the classroom? *Scan: The Journal for Educators, 41*(6), 8–16. https://doi.org/10.3316/informit.627739134125076

Checa, D., & Bustillo, A. (2020). A review of immersive virtual reality serious games to enhance learning and training. *Multimedia Tools and Applications, 79*(9), 5501–5527. https://doi.org/10.1007/s11042-019-08348-9

Christopoulos, A., Conrad, M., & Shukla, M. (2018). Increasing student engagement through virtual interactions: How? *Virtual Reality, 22*, 353–369. https://doi.org/10.1007/s10055-017-0330-3

Collie, R. J., Martin, A. J., Bobis, J., Way, J., & Anderson, J. (2018). How students switch on and switch off in mathematics: Exploring patterns and predictors of (dis)engagement across middle school and high school. *Educational Psychology, 39*(4), 489–509. https://doi.org/10.1080/01443410.2018.1537480

Cooper, E. (2010). Tutoring Center Effectiveness: The Effect of Drop-In Tutoring. *Journal of College Reading and Learning, 40*(2), 21–34. https://doi.org/10.1080/10790195.2010.10850328

Csikszentmihalyi, M., & Csikszentmihalyi, I. (Eds.). (1988). *Optimal experience*. Cambridge University Press.

Dirkx, K. J. H., Camp, G., Kester, L., & Kirschner, P. A. (2019). Do secondary school students make use of effective study strategies when they study on their own? *Applied Cognitive Psychology, 33*(5), 952–957. https://doi.org/10.1002/acp.3584

Doyle, N. E., McDowall, A., Randall, R., & Knight, K. (2022). Does it work? Using a Meta-Impact score to examine global effects in quasi-experimental intervention studies. *PloS one*, *17*(3), Article e0265312. https://doi.org/10.1371/journal.pone.0265312

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motivation* (pp. 75–146). W. H. Freeman.

Eccles, J. S. (2013). Gender and STEM: Opting in versus dropping out. *International Journal of Gender, Science and Technology, 5*(3), 184–186. https://genderandset.open.ac.uk/index.php/genderandset/article/view/315/529

Field, A. P., Evans, D., Bloniewski, T., & Kovas, Y. (2019). Predicting maths anxiety from mathematical achievement across the transition from primary to secondary education. *Royal Society Open Science, 6*, 1–13. https://doi.org/10.1098/rsos.191459

Fokides, E. (2018). Digital educational games and mathematics. Results of a case study in primary school settings. *Education and Information Technologies*, *23*(2), 851–867. https://doi.org/10.1007/s10639-017-9639-5

Fong, C. J., Kremer, K. P., Cox, C. H. T., & Lawson, C. A. (2021). Expectancy-value profiles in math and science: A person-centered approach to cross-domain motivation with academic and STEM-related outcomes. *Contemporary Educational Psychology*, *65*, Article 101962. https://doi.org/10.1016/j.cedpsych.2021.101962

Fritz, C. O., Morris, P. E., Acton, M., Voelkel, A. R., & Etkind, R. (2007). Comparing and combining retrieval practice and the keyword mnemonic for foreign vocabulary learning. *Applied Cognitive Psychology, 21*, 499–526. https://doi.org/10.1002/acp.1287

Ghasemi, E., & Burley, H. (2019). Gender, affect, and math: A cross-national meta-analysis of Trends in International Mathematics and Science Study 2015 outcomes. *Large-Scale Assessments in Education, 7*(1), 1–25. https://doi.org/10.1186/s40536-019-0078-1

Girvan, C. (2018). What is a virtual world? Definition and classification. *Educational Technology Research and Development, 66*(5), 1087–1100. https://doi.org/10.1007/s11423-018-9577-y

Harun, Kartowagiran, B., & Manaf, A. (2021). Student attitude and mathematics learning success: A metaanalysis. *International Journal of Instruction*, *14*(4), 209–222. https://doi.org/10.29333/iji.2021.14413a

Hookham, G., & Nesbitt, K. (2019). A systematic review of the definition and measurement of engagement in serious games. In *Proceedings of the Australasian Computer Science Week Multiconference* (pp. 1–10). Association for Computing Machinery. https://doi.org/10.1145/3290688.3290747

Hulse, T., Daigle, M., Manzo, D., Braith, L., Harrison, A., & Ottmar, E. (2019). From here to there! Elementary: A game-based approach to developing number sense and early algebraic understanding. *Educational Technology Research and Development, 67*(2), 423–441. https://doi.org/10.1007/s11423-019-09653-8

Hwang, Y., Shin, D., & Lee, H. (2023). Students' perception on immersive learning through 2D and 3D metaverse platforms. *Educational Technology Research and Development, 71*, 1687–1708. https://doi.org/10.1007/s11423-023-10238-9

Jakaitienė, A., Želvys, R., Vaitekaitis, J., Raižienė, S., & Dukynaitė, R. (2021). Centralised mathematics assessments of Lithuanian secondary school students: population analysis. *Informatics in Education, 20*(3), 439–462. https://doi.org/10.15388/infedu.2021.18

Jing, T. W., Yue, W. S., & Murugesan, R. K. (2015). Learning outcome enhancement via serious game: Implementing game-based learning framework in blended learning environment. In *2015 5th International Conference on IT Convergence and Security (ICITCS)* (pp. 1–3). IEEE. https://ieeexplore.ieee.org/abstract/document/7292992

Kalogiannakis, M., Papadakis, S., & Zourmpakis, A. I. (2021). Gamification in science education. A systematic review of the literature. *Education Sciences*, *11*(1), 22. https://doi.org/10.3390/educsci11010022

Karpicke, J. D., Butler, A. C., & Roediger, H. L., III. (2009). Metacognitive strategies in student learning: Do students practice retrieval when they study on their own? *Memory, 17*, 471–479. https://doi.org/10.1080/09658210802647009

Khaleel, F. L., Noraidah, S., Tengku, S. M. T. W., & Amirah, I. (2016). The architecture of dynamic gamification elements based learning content. *Journal of Convergence Information Technology, 11*(3), 164–177.

Khan, A., Ahmad, F. H., & Malik, M. M. (2017). Use of digital game based learning and gamification in secondary school science: The effect on student engagement, learning and gender difference. *Education and Information Technologies, 22*(6), 2767–2804. https://doi.org/10.1007/s10639-017-9622-1

Kim, H. Y. (2013). Statistical notes for clinical researchers: Assessing normal distribution (2) using skewness and kurtosis. *Restorative Dentistry & Endodontics, 38*(1), 52–54. https://doi.org/10.5395/rde.2013.38.1.52

Klee, H. L., & Miller, A. D. (2019). Moving Up! Or Down? Mathematics Anxiety in the Transition from Elementary School to Junior High. *The Journal of Early Adolescence, 39*(9), 1–26. https://doi.org/10.1177/0272431618825358

Krath, J., Schürmann, L., & Von Korflesch, H. F. (2021). Revealing the theoretical basis of gamification: A systematic review and analysis of theory in research on gamification, serious games and game-based learning. *Computers in Human Behavior, 125*, 106963. https://doi.org/10.1016/j.chb.2021.106963

Laine, T. H., & Lindberg, R. S. (2020). Designing engaging games for education: A systematic literature review on game motivators and design principles. *IEEE Transactions on Learning Technologies, 13*(4), 804–821. https://doi.org/10.1109/TLT.2020.3018503

Lei, H., Chiu, M. M., Wang, D., Wang, C., & Xie, T. (2022). Effects of game-based learning on students' achievement in science: A meta-analysis. *Journal of Educational Computing Research, 60*(6), 1373–1398. https://doi.org/10.1177/07356331211064543

Li, Y., & Schoenfeld, A. H. (2019). Problematizing teaching and learning mathematics as "given" in STEM education. *International Journal of STEM Education, 6*(1), 1–13. https://doi.org/10.1186/s40594-019-0197-9

Marto, A., & Gonçalves, A. (2022). Augmented reality games and presence: A systematic review. *Journal of Imaging, 8*(4), 91. https://doi.org/10.3390/jimaging8040091

Michaelides, M. P., Brown, G. T. L., Eklöf, H., & Papanastasiou, E. C. (2019). *Motivational profiles in TIMSS mathematics: Exploring student clusters across countries and time* (Vol. 7). Springer. https://doi.org/10.1007/978-3-030-26183-2_2

Michailidis, L., Balaguer-Ballester, E., & He, X. (2018). Flow and immersion in video games: The aftermath of a conceptual challenge. *Frontiers in Psychology*, *9*, Article 1682. https://doi.org/10.3389/fpsyg.2018.01682

Mount, N. J., Chambers, C., Weaver, D., & Priestnall, G. (2009). Learner immersion engagement in the 3D virtual world: principles emerging from the DELVE project. *Innovation in Teaching and Learning in Information and Computer Sciences, 8*(3), 40–55. https://doi.org/10.11120/ital.2009.08030040

Moyer-Packenham, P. S., Lommatsch, C. W., Litster, K., Ashby, J., Bullock, E. K., Roxburgh, A. L., Shumway, J. F., Speed, E., Covington, B., Hartmann, C., Clarke-Midura, J., Skaria, J., Westenskow, A., MacDonald, B., Symanzik, J., & Jordan, K. (2019). How design features in digital math games support learning and mathematics connections. *Computers in Human Behavior*, *91*, 316–332. https://doi.org/10.1016/j.chb.2018.09.036

Naik, N. (2015). The use of GBL to teach mathematics in higher education. *Innovations in Education and Teaching International, 54*(3), 238–246. https://doi.org/10.1080/14703297.2015.1108857

Nakamura, J., & Csikszentmihalyi, M. (2014). The concept of flow. In M. Csikszentmihalyi (Ed.), *Flow and the Foundations of Positive Psychology* (pp. 239–263). Springer. https://doi.org/10.1007/978-94-017-9088-8_16

Nuthall, G. (1999). The way students learn: Acquiring knowledge from an integrated science and social studies unit. *The Elementary School Journal, 99*(4), 303–341. https://doi.org/10.1086/461928

Nuthall, G. (2005). The cultural myths and realities of classroom teaching and learning: A personal journey. *Teachers College Record, 107*(5), 895–934. https://doi.org/10.1111/j.1467-9620.2005.00498.x

Nuthall, G. (2007). *The hidden lives of learners*. NZCER Press.

Nuthall, G. (2004). Relating classroom teaching to student learning: A critical analysis of why research has failed to bridge the theory-practice gap. *Harvard Educational Review*, *74*(3), 273–306. https://doi.org/10.17763/haer.74.3.e08k1276713824u5

Organisation for Economic Cooperation and Development. (2013). *PISA 2012 results: Ready to learn: Students' engagement, drive and self-beliefs* (Vol. III). https://doi.org/10.1787/9789264201170-en

Özhan, Ş. Ç., & Kocadere, S. A. (2019). The effects of flow, emotional engagement, and motivation on success in a gamified online learning environment. *Journal of Educational Computing Research, 57*(8), 2006–2031. https://doi.org/10.1177/0735633118823159

Semeraro, C., Giofrè, D., Coppola, G., Lucangeli, D., & Cassibba, R. (2020). The role of cognitive and non-cognitive factors in mathematics achievement: The importance of the quality of the student-teacher relationship in middle school. *Plos One*, *15*(4), Article e0231381. https://doi.org/10.1371/journal.pone.0231381

Sun, Z., Xie, K., & Anderman, L. H. (2018). The role of self-regulated learning in students' success in flipped undergraduate math courses. *The Internet and Higher Education, 36*, 41–53. https://doi.org/10.1016/j.iheduc.2017.09.003

Tcha-Tokey, K., Christmann, O., Loup-Escande, E., & Richir, S. (2016). Proposition and validation of a questionnaire to measure the user experience in immersive virtual environments. *International Journal of Virtual Reality, 16*(1), 33–48. https://doi.org/10.20870/IJVR.2016.16.1.2880

Tran, L. T., & Nguyen, T. S. (2021). Motivation and mathematics achievement: A Vietnamese case study. *Journal on Mathematics Education, 12*(3), 449–468. https://doi.org/10.22342/jme.12.3.14274.449-468

Tsai, Y. L., & Tsai, C. C. (2020). A meta-analysis of research on digital game-based science learning. *Journal of Computer Assisted Learning*, *36*(3), 280–294. https://publons.com/publon/10.1111/jcal.12430

UK National Recognition Information Centre. (2004). *Grading transfer system increasing transparency of access qualifications for higher education in Europe*. Retrieved Dec. 23, 2023 from https://www.ecctis.com/documents/contributions/Grading-Transfer-System-Methodology.pdf

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wang, S. M., Chen, Y. C., Hou, H. T., Hsu, H. Y., & Li, C. T. (2020). Exploring the effects of card gamebased gamification instructional activity on learners' flow experience, learning anxiety, and performance-A preliminary study. In So, H. J. et al. (Eds.) *ICCE 2020 28th International Conference on Computers in Education* (Vol. 2, pp. 190–198). https://apsce.net/icce/icce2020/proceedings/W1-13/W4/ICCE2020-Proceedings-Vol2-W4_5.pdf

Xu, J. (2020). Longitudinal effects of homework expectancy, value, effort, and achievement: An empirical investigation. *International Journal of Educational Research*, *99*, Article 101507. https://doi.org/10.1016/j.ijer.2019.101507

Yadav, S. (2019). Role of mathematics in the development of society. *IJRAR-International Journal of Research and Analytical Reviews*, 6(4), 295–298. https://ssrn.com/abstract=3677993

Authors and Affiliations

Athanasios Christopoulos[1] · Stylianos Mystakidis[2,3] · Justyna Kurczaba[1] · Mikko-Jussi Laakso[1,4] · Chrysostomos Stylios[5,6]

✉ Athanasios Christopoulos
atchri@utu.fi

1 Turku Research Institute of Learning Analytics, Faculty of Science, University of Turku, Turku, Finland

2 School of Natural Sciences, University of Patras, Patras, Greece

3 School of Humanities, Hellenic Open University, Patras, Greece

4 Department of Computing, Faculty of Technology, University of Turku, Turku, Finland

5 Department of Informatics and Telecommunications, University of Ioannina, Arta, Greece

6 Industrial Systems Institute, Athena Research Center, Patras Science Park Building, Platani, Patras, Greece