2023

Gamified Learning in Math Education

Nathan Michael Voigt
Bethel University

GAMIFIED LEARNING IN MATH EDUCATION

A MASTER'S THESIS

SUBMITTED TO THE FACULTY

OF BETHEL UNIVERSITY

BY

NATHAN VOIGT

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF ARTS

May 2023

BETHEL UNIVERSITY

GAMIFIED LEARNING IN MATH EDUCATION

Nathan Voigt

May 2023

APPROVED

Advisor's Name: Meghan Cavalier, Ed.D.
Program Director's Name: Katie Bonawitz, Ed.D.

Acknowledgements

I would like to thank Dr. Cavalier for guiding me through the thesis writing process and Dr. Wickam for administrating the 5-12 licensure program.

Abstract

Educators are always looking for ways to improve engagement in the classroom. Traditional approaches to math education are being phased out, amid declining student achievement in mathematics, and replaced with innovations like gamified learning. Gamified learning supplements instruction and curriculum by providing a fun environment, ensconced in virtual and augmented reality, for students to interact with the content. The guiding research question of this thesis is focused on the efficacy of gamified learning in mathematics with respect to motivation and student achievement. Studies in which researchers created and tested new math games have been evaluated to answer the guiding research question. Self-determination Theory, a theory about human motivation, is unpacked and used as a framework that helps educators generalize the conclusions to the classroom. The results showed gamified learning to be a competent tool for increasing motivation and student achievement in the short term, but more longitudinal studies are needed to verify such gains can be sustained over the span of a course.

Table of Contents

CHAPTER ONE: INTRODUCTION

Introduction

Is the decline in math proficiency due to a lack of student motivation? Or lack of innovation in the math classroom? A recent survey conducted by Global Strategy Group and funded by the Bill and Melinda Gates Foundation found that most parents and teachers are unsatisfied with how Kindergarten-12 math education is taught in U.S. schools, with only 50% believing learning math is currently engaging (Jimenez, 2023). A wave of change is sweeping over math classrooms around the United States as educators seek to replace the traditional classroom model with a more engaging experience. Yet educators ought to mindfully balance the benefits with the costs when implementing any innovation. Change is expensive, and so are the resources needed to implement the modern innovations in question in the context of large-scale studies (Shi et al., 2022). Still, technology is at the core of gamified learning, making gamified learning compatible with other technological innovations. Gamified learning is one potential innovation educators can use to support the instruction and content in the math classroom (Landers, 2014).

According to Zainuddin (2018), gamified learning is a combination of gamification and classroom instruction that can potentially augment students' learning capabilities. Gamified learning generally features a student-centered learning experience and involves using technology in nuanced ways; however, certain games may require a technological overhaul of the course, while others can be streamlined into preexisting units more easily. Implementing high quality instructional content or using other technology in the classroom is independent of the game being played and beyond the scope of this thesis; however, the implementation of gamified learning cannot affect outcomes on achievement and motivation without the presence of high quality instruction (Landers, 2014). In addition to having a symbiotic relationship with instructional content, gamified learning can be paired with different classroom models, such as a flipped classroom as observed in the study conducted by Zhao et al. (2021).

Different mediums, elements, and mechanics in games offer the user a unique experience. Mediums, such as augmented reality and virtual reality make up an engaging environment for students to explore. Mechanics are features in a game that students can manipulate like touch-screen compatibility (Hulse et al., 2019). According to Park and Kim (2021) badges and leaderboards are the most common elements in gamified learning. Leaderboards display the progress of a subset of all participants and contribute to the social learning of students through collaboration (Jagust et al., 2018; Shanmugam & Mohamed, 2020 in Rincon-Flores et al., 2023) or competition (Jagust et al., 2018). Badges can be earned by completing tasks and are visual reminders of success (Rincon-Flores et al., 2023). Games have many other elements and mechanics that will be highlighted in Chapter Two. Each element could have causal effects on student motivation and achievement or the games could be greater than the sum of their parts. Some researchers choose to isolate elements to limit confounding variables and others choose to package everything together.

Rationale

The current research on gamified learning reviewed in this thesis based conclusions about student achievement on scores from tests or evaluations created for the study, rather than on classroom test scores, because few long-term studies exist. Data from self-report questionnaires taken by students were used to form conclusions on gains in motivation. Though student achievement can be considered a function of motivation (Steinmayr et al., 2019), the researchers in this literature review often discussed the results of the studies with respect to motivation independently. The student learning experience is also of importance because this affects how students view math. If Kindergarten-12 students do not have a positive experience learning math, they may not continue to pursue higher-level math courses or be motivated to learn more. This thesis will answer the guided research question by exploring various studies, which have briefly implemented gamified learning in K-12 math classrooms. The mediums researchers use to answer their research questions generally boil down to specially crafted assessments and self-report surveys. Though most studies analyzed in this thesis base their conclusions on experiments conducted on students, Rincon Flores et al. (2023) covered professors' observations of their students playing games. The concepts covered in the games usually had not been learned formally before in class.

Though gamified learning is possible in all content areas, for the scope of this thesis, it will be discussed in the context of Kindergarten-12 math classrooms. In Chapter Two, Self-Determination Theory, general gamified learning, and gamified learning in math classrooms will be discussed. More specifically, the research questions, methodology, and conclusions of nine research studies relevant to gamified learning in math education are laid out in Chapter Two.

Definitions of Terms

Academic Motivation Scale- a 28 item test, scored on a Likert scale, that measures educational motivation (Vallerand et al., 1992).

Augmented Reality- virtual images that are superimposed over real-world images (Azuma, 1997).

Game-Based Virtual Reality Learning Environment- is a computer generated, three-dimensional world that uses immersive technology to make users feel like they are present (Shi et al., 2022).

Gamification- the process of adding game elements to non-gaming content (Deterding et al., 2011).

HEXAD Gamification User Typology - is a model that represents six player profiles that outline different motivations for playing games (Marczewski, 2015).

HEXAD Questionnaire- a survey that is scored to determine the correct player profile from the six profiles in the HEXAD model (Marczewski, 2015).

Needs - According to Self-Determination Theory, needs are innate psychological nourishments that provides psychological growth, prosperity, and integrity (e.g., autonomy, competence, and relatedness) (Deci & Ryan, 2000).

Self-Report Bias- "a methodological problem that arises when researchers rely on asking people to describe their thoughts, feelings, or behaviors rather than measuring these directly and objectively. People may not give answers that are fully correct, either because they do not know the full answer or because they seek to make a good impression." (American Psychological Association, n.d.).

Statement of Guiding Research Question

This literature review will focus on exploring the potential of gamified learning in math classrooms and its subsequent impact on student achievement and motivation. The guiding research question is: how does gamified learning increase motivation and achievement in the math classroom?

CHAPTER II: LITERATURE REVIEW

Self-Determination Theory

Self-Determination Theory (SDT), proposed by Deci and Ryan (2000), is an approach to understanding human motivation and behavior, derived explicitly from evolved, intrapersonal resources related to regulating intrinsic and extrinsic motivations. Students engaged in activities because of a grade, future career, or pleasing their guardians are all extrinsically motivated, whereas students who are genuinely interested in the content are intrinsically motivated to participate (Ryan & Deci, 2000). Self-Determination Theory strengthens the conclusions of some research studies presented in this thesis and provides a baseline for best practices in cultivating motivation and, consequently, student achievement. Ryan and Deci (2000) introduced autonomy, competence, and relatedness as three basic psychological needs for enhancing learning motivation. Autonomy is the human need to make unforced decisions. Competence refers to the desire to be effective in one's work. Relatedness is the need to be connected to others in a community (Ryan & Deci, 2000). Therefore, a model aligned with SDT allows students to engage in desirable activities insofar as they want to work towards mastery of the content without being isolated from their peers.

Ryan and Deci (2000) posited that humans are active organisms that are intrinsically motivated to do challenging, novel, and interesting tasks. Intrinsic motivation allows humans to be actively engaged in tasks they want to perform; however, for this engagement to be sustained, psychological needs, namely autonomy and competence, must be satisfied. Generally, intrinsic motivation flourishes under conditions when autonomy, competence, and relatedness are satisfied. On the other hand, intrinsic motivation is undermined when these psychological needs have not been met. For example, intrinsic motivation is undermined when rewards have been introduced because people may ignore their preference and subject themself to the more rewarding task instead (Ryan & Deci, 2000).

Feedback is related to competence. According to Ryan and Deci (2000), when people received positive feedback, their ability to succeed was validated in a particular task and intrinsic motivation for that task increased. On the other hand, negative feedback can decrease intrinsic motivation. Optimal conditions for fostering intrinsic motivation are created when autonomy and competence are satisfied (Ryan & Deci, 2000).

Creating conditions where autonomy, competence, and relatedness are satisfied is foundational to increasing motivation; however, the satisfaction of relatedness is not always required to increase intrinsic motivation because humans may be intrinsically motivated in many activities in which they choose to participate alone (Ryan & Deci, 2000). Satisfying relatedness is not even possible with all activities (Ryan & Deci, 2000). Ryan and Deci (2000) believed humans are able to develop a "secure relational base" (p. 9), are more likely to experience intrinsic motivation, and are more resilient to future events that decrease intrinsic motivation. In an academic setting, the connections students make with their teachers and peers matter. Intrinsic motivation can be increased when teachers are perceived as caring (Ryan & Grolnick, 1986; Ryan et al., 1994).

Ideally, students are motivated to engage in activities freely because they genuinely enjoy learning, but many activities in school require self-regulation; they are not interesting enough to be fueled by intrinsic motivation (Deci et al., 1994). Deci and Ryan (2000) defined extrinsic motivation as the motivation to perform a certain activity in order to attain outcomes indirectly related to the activity. Extrinsic motivation appears with varying degrees of autonomy, but generally, autonomous learning is associated with intrinsic motivation (Deci & Ryan, 2000).

Thus, if autonomy is satisfied by the game, students who are intrinsically motivated may experience gains in motivation. However, games that give students options may help extrinsically motivated students become autonomously motivated or even intrinsically motivated too. Deci and Ryan (2000) said, "As people internalize regulations and assimilate them to the self, they experience greater autonomy in action" (p. 6). Moreover, activities that support autonomy, competence, and relatedness cultivated more robust levels of extrinsic motivation (Deci & Ryan, 2000). In fact, as the level of autonomy increased in extrinsic motivation, so did student engagement and achievement (Connell & Wellborn, 1991; Miserandino, 1996). When drawing conclusions from research studies, it is impossible to ascertain the degree to which students have assimilated their new responsibilities, only whether a given game offers students an embedded system of choices. Giving students choices models the process of internally regulating their learning and opens the door for autonomous motivation in the future (Deci & Ryan, 2000).

Self-Determination Theory remains relevant in an educational setting (Niemiec & Ryan, 2009). According to Niemiec and Ryan (2009), when students were inevitably faced with uninteresting tasks in school, intrinsic motivation took a back seat to extrinsic motivation. But not all forms of extrinsic motivation are of equal potency. Students who have internalized extrinsic motivation exhibited more autonomous types of extrinsic motivation, which in turn leads to an augmented learning experience. Furthermore, psychological need satisfaction facilitated internalization (Niemiec & Ryan, 2009). Games that cultivated need satisfaction augmented motivation in students who may not otherwise be intrinsically motivated. Many studies showed that intrinsic motivation and autonomous extrinsic motivation impacted student engagement and academic achievement positively (Niemiec & Ryan, 2009). Taylor et al. (2014)

conducted a meta-analysis of longitudinal and cross-sectional studies that covered a variety of cultural and academic settings, and intrinsic motivation was found to be the most impactful of any motivation on student achievement.

Consequently, the impact of gamification on motivation will be evaluated with these needs in mind. Applying need satisfaction analysis to some of the research studies in this literature review will further corroborate the conclusions of these research studies and make generalized conclusions more accurate. Concluding whether intrinsic or extrinsic motivation impacted student learning and achievement more is beyond the scope of this literature review. The subtlety between more autonomous forms of extrinsic motivation and intrinsic motivation made empirically discerning the difference problematic (Flunger et al., 2022). Rather, impacts on motivation will be taken as a whole.

General Gamified Learning

The gamification of learning is adding game elements to non-gaming subject matter in the classroom to supplement student learning (Deterding et al., 2011). The basic premise behind gamification in education is that it will provide an immersive experience, similar to that of playing games, that motivates and inspires the user to keep learning (Codish & Ravid, 2014; McGonigal, 2011). Gamification was not a widely known term until 2010; however, there is nothing new about incentivizing learning (Dichev & Dicheva, 2017). Students learn by playing games set in augmented reality (AR), which presents information to students in an interactive environment containing a combination of 3-D, real-world, and virtual elements (Azuma T., 1997). According to Seaborn and Fels (2015), gamification, in the context of education, has received increased attention from educators. Of course, criticism follows an increase in popularity. Extrinsic motivation does not motivate students to keep playing the game

(Sánchez-Franco, 2009 as cited in Mahnic, 2014). The primary motivation elicited by playing games using rewards-based elements, such as points and badges, is extrinsic motivation (Mahnic, 2014). While studies incorporating reward-based elements will be analyzed, more robust games that include narrative, touch-screen intractability, and a virtual reality world will also be unpacked.

Elements, Mechanics, and Dynamics of Gamification

Gamification has been difficult to evaluate because of a lack of uniform elements, and educators and researchers could benefit from more centralized guidance regarding game design in general (Klabbers, 2018). Even so, common elements can be found throughout various games. Not all elements are equal because some are derived from systemic game design, while others, such as points, badges, and leaderboards, are more superficial (Manzano-León et al., 2021). Manzano-León et al. (2021) conducted a meta-analysis of 14 studies, selected from 198 studies by a thorough set of inclusion criteria, that counted the number of studies that had each element from this list of mechanics: rewards, prize, achievements, points, levels, badges, and ranking. Game dynamics provide the context for the elements and consist of: challenge, playful activities, events, tasks, roles, feedback, choices, and competition. Independently existing with the elements and dynamics, an overarching narrative is becoming more widespread; more than half of these games incorporated a narrative (Manzano-León et al., 2021). By compiling a list of all elements and dynamics occurring in the studies reviewed in the interdisciplinary meta-analysis, an implicit argument is made for their commonality across all games and is corroborated by the additional studies reviewed in this paper.

Theory of Gamification

Landers (2014) proposed a theory of gamification to help expand the cannon of gamification and guide researchers and educators alike in game design. The theory states that gamification mediates and moderates behaviors or attitudes that affect learning outcomes. The theory is rooted in two self-evident postulates, namely, “Instructional content influences learning outcomes and behaviors” and “behaviors/attitudes influence learning” (p. 760). Gamification can *mediate* the causal relationship between instruction content and learning outcome by providing an engaging environment where students spend more time interacting with the material. Thus, the resulting learning outcome cannot solely be caused by quality instruction, though the learning outcome may be irrelevant if the instruction and curriculum lack quality. Gamification *moderates* outcomes by supplementing and strengthening instructional content, a primary causal agent of student motivation and achievement (Landers, 2014). According to Caponetto et al. (2014), gamification is capable of supporting a variety of behaviors, such as retention of content, collaboration with peers, and assignment completion. For game elements to be effective, with respect to producing a given learning outcome, such as finishing a practice test, the prerequisite behaviors of fulfilling the learning outcome must be linked to increased learning, and the game elements must cause these behaviors (Landers, 2014). Essentially, a game is only as good as the instruction and curriculum it is associated with.

While much of the science behind creating motivating mechanics and elements is beyond the scope of this paper, some game elements, which are largely interchangeable and solely undefining, will be isolated and discussed within the context of their efficacy on motivation and achievement when possible. Unfortunately, many current research studies present a game as a complete package with a plethora of different elements for their experiment, resulting in a conflation of casual elements (Landers, 2014). More research must be done where elements are implemented separately to help distinguish effective elements from unnecessary or detrimental elements.

Gender Bias

Much has been said about the benefits of gamified learning, but unfortunately, gamified learning can unwittingly expose gender biases. Stereotype threat is a phenomenon that occurs when someone has an opportunity to validate known stereotypes about a demographic they are a member (Steele, 1998). Researchers like Albuquerque et al. (2017), want to know how stereotype threat can manifest itself in a gamified setting and its ramifications on anxiety. For the purpose of this study, Albuquerque et al. (2017) created three variations of a gamified-tutoring application, namely stereotyped-male, stereotype-female, and non-stereotyped. One hundred and twenty-seven Brazilians (85 male and 45 female) volunteered to participate in the study. They came from a wider age range than school-age students (18-54), but most were in the 18-34 age range (Albuquerque et al., 2017). The wider age range should not be alarming, but more studies involving school-age children interacting with gamified learning would be helpful in generalizing results. Albuquerque et al. (2017) had them take a pre and post anxiety test, which consisted of 20, Likert scale questions, in between using the gamified-tutoring app, which consisted of 60 multidisciplinary questions. In the gamified app, which all groups use, participants can gain or lose points, earn badges, and progress to higher levels. The gamified tutoring app was differentiated for the gender stereotype groups. The layout and customization options were tailored to create male and female themes; however, both themes were used by male and female participants (Albuquerque et al., 2017).

After statistical analysis, Albuquerque et al. (2017) found that there are differences in anxiety levels among the men and women participants. Women who used the male-stereotyped app had significant differences in anxiety levels compared to those who participated in the non-stereotyped and female-stereotyped apps. Admittedly, the gamified environment explored in this study may not be a good match for other games, and thus, the results should be generalized with caution. Even the presence of gamification elements in the tutoring app, such as a point and leaderboard system, could be a source of anxiety among gender minority participants (Albuquerque et al., 2017). Far from an indictment of gamified learning, the results warn educators that certain combinations of game elements can increase anxiety because games are composed of unique elements. Yet these results point other researchers in the right direction; away from implementing too many non-academic game elements and towards embedding pedagogically sound elements, like touch-screen interactivity, into their games.

Gamified Learning in Math

The gamification of math classrooms is an important aspect of many math classrooms around the United States. Students enjoy practicing math concepts in a fun and interactive setting. Consequently, researchers work to infuse mobile and computer games with math concepts in hopes of providing students with the motivation that this kind of entertainment provides. Different versions of gamified models will be discussed in the following sections, as researchers work to streamline and standardize gamification in math.

The world for many of these games is made possible by advances in technology, such as augmented reality, which can display graphs, real-world objects and environments, and virtual information with computer-generated technology (Li et al., 2022). Ahmad and Jainani (2022) found important benefits of augmented reality after reviewing 19 journal articles between 2015 and 2019. Moreover, environments created by augmented reality increase self-confidence and mathematical understanding, enhance visualization skills, and promote interactive learning (Ahmad & Jainani, 2022).

Interactive Augmented Reality

Some games are set in augmented reality (AR), which presents information to students in an interactive environment, containing a combination of three-dimensional, real-world, and virtual elements (Azuma, 1997). Augmented reality can be thought of as a derivative of virtual reality because AR virtual elements are used in tandem with real-world objects (Azuma, 1997 in Sirakaya & Sirakaya, 2022). According to Sirakaya and Sirakaya (2022) AR can be used to bring a myriad of math concepts to life in rich environments. Over time, AR has become less expensive and easier to develop on computers and phones (Sirakaya & Sirakaya, 2022).

In one such study, Rebollo et al. (2022) integrated multiplication facts with mobile games by designing two turn-based mini-games in which students performed multiplication operations to win. Results supported the conclusion that the mobile games created for the study were effective in engaging students in the learning of multiplication tables (Rebollo et al., 2022).

Researchers like Rebollo et al. (2022) are interested in how gamification through augmented reality affects student learning in mathematics classrooms. Rebollo et al. (2022) developed two interactive mobile games which help students learn their multiplication tables. The study involved 37 students ranging from eight to nine years old. The control group practiced their multiplication tables traditionally, while the experimental group practiced them by playing the games. It is important to note that all students participating in the study had previously been introduced to multiplication tables in school. Students completed their problems at home with their guardians' consent. So, the games are meant to support student learning through practice and help consolidate their knowledge. In the game, *Battle Against the Colossus*, students answer multiplication questions to defeat the mighty colossus. When a question is answered correctly the boss' health decreases. Similarly, the user seeks to defend their life by answering multiplication questions to thwart the boss' attacks. The user wins when the health bar of Colossus is reduced to zero or the user has more health after 30 multiplication questions are attempted. Another game, *Save the Planet*, involves students throwing cans into the recycling bin by answering multiplication questions correctly; whereas incorrect answers lead to the planet slowly destroying itself, observed through visual effects. It is important to note that students must answer questions within 10 seconds (Rebollo et al., 2022). Both games are limited, in their current form, to a multiple-choice format; neither game has the capacity to help students practice concepts that would require multiple steps.

The hypothesis of Rebollo et al. (2022) was two-fold. First, student learning of multiplication tables will be positively impacted by the practice provided by each game. Second, this AR provides a more engaging atmosphere than traditional methods of learning multiplication tables. To assess the learning effectiveness of the games and traditional learning, each group of students took a pretest in which they were asked self-assessing questions about their current knowledge of multiplication tables and to perform multiplication exercises. Then, after the students played the game at home a half an hour a day for one week, the students in the experimental group were given a posttest consisting of multiplication questions and questions pertaining to engagement and usability of the game (Rebollo et al., 2022).

The learning effectiveness of each method was determined by comparing the pretest and posttest results of the multiplication questions students were asked to perform. In both groups, correct answers increased, so the hypothesis that student learning was positively impacted by playing the game is confirmed. The second hypothesis was confirmed based on the results from the 13-question usability (engagement) questionnaire, in Likert scale form, given to the control group. They found that 94% of these students said that it was easy to answer the multiplication questions in 10 seconds and that they were not distracted by learning through a video game. The usability results imply that the games were engaging to students. In summary, playing the game positively impacted the learning and engagement of students (Rebollo et al., 2022). Therefore, learning multiplication tables through games is a valid replacement for learning multiplication tables through traditional methods.

Li et al. (2022) explored the verbal and real-world capabilities of AR in the context of supporting student learning of linear functions. They wanted to know how students perceive and interact with the AR environment to support the learning of linear functions through multiple representations. They worked with two middle school classes in Shandong, China. Each student was given a pretest from which they were grouped by ability; researchers were able to observe how each group interacted with the AR (Li et al., 2022).

Li et al. (2022) focused on the application of linear functions to uniform linear motion. They isolated three relevant characteristics of uniform linear motion (mapping, covariance, and function as object) for students to match to three representations of these characteristics (real-life, symbolic, and graphical). The high-achieving ability group reported that learning in AR helped them have a better understanding of linear functions. The low-achieving students liked AR because it was simple to use and helped them visualize aspects of linear functions. Both ability groups said AR was interesting. When designing the AR for linear functions, the researchers recommend designing apps that scaffold the progression of students, because students may choose an easier AR game over more difficult ones (Li et al., 2022). Clearly, interesting learning tasks are more engaging than disinteresting tasks; interesting tasks may be the bridge that low-achieving students need to reach higher achievement levels.

Narrative

Gamified Learning with a Flipped Classroom

Some educators choose to blend gamified learning with their flipped classroom in order to offer a more guided learning experience for certain topics. Zhao et al. (2021) conducted their research on a primary school in Chengdu China. They divided 130 students into three equal groups to study the effects of combining gamified and flipped classroom elements. Two groups were used to control for the benefits of traditional instruction and normal flipped classroom learning. Specifically, they wanted to know whether gamified learning in a flipped classroom setting yielded the highest student achievement, motivation, and meta-cognition when compared to a flipped classroom and a traditional classroom. After two weeks, students completed a written test pertaining to what they learned in their lessons about fractions and a self-report questionnaire about motivation. All students had the same prior knowledge and had no previous experience with the new content about fractions (Zhao et al., 2021). Before moving into a discussion of the results, it is important to provide a context to such conclusions. What does this particular gamified learning experience entail? Are the gaming elements of the interactive e-book easy to streamline with a flipped classroom learning experience?

Interactive Learning E-Book

The interactive learning e-book Zhao et al. (2021) created consisted of three modules, namely, learning, practice, and gamified learning. The modules are as straightforward as they seem. In the first module, students learn the new content through a story set in a snowy castle somewhere in Alaska. They set out as a warrior tasked with uncovering the mystery of the bright colors emitted from the castle at midnight. It turns out that each color corresponds to an elf waiting for students to help them solve a mystery. When students find the elves, they are introduced to the fractional concepts they need to solve problems later on. After students are given time to practice the concepts within the gamified learning environment, they take a quiz at the end to earn a badge. If they do not earn a satisfactory badge, they may look over the feedback given and re-attempt the questions they scored poorly on before. The last step of awarding badges based on the number of questions answered correctly was classified as the gamified step, but all three phases can be safely classified as gamified learning. Throughout this process, students were able to interact with the material by moving pictures or clicking on pictures (Zhao et al., 2021). Given the pictures shared of the interactive e-book, it is plausible that this environment would be decidedly interesting to a third grader.

Zhao et al. (2021) bolstered the gamified e-book with Self-Determination Theory (SDT) by designing it to satisfy the psychological needs of autonomy, competence, and relatedness. Again, when conditions are present to cultivate these needs, student motivation ought to increase. Students were able to explore the environment of the game and take control of their own learning. If they failed to earn the badge they wanted, they could go back and re-learn the content to earn the badge they wanted. For the purposes of this study, students were regulated to one of the three groups, but under normal conditions, teachers could give students the option of playing the game to learn fractions versus learning them through a video. Next, competence is fostered through practice problems designed with appropriate difficulty levels to scaffold the new material and prepare them for the quiz. Students may feel accomplished when they earn their badge at the end. Finally, students were able to share the results with their classmates, precipitating social interaction when students discussed the game elements with their friends

(Zhao et al., 2021). While it is certain that students using the interactive e-book were able to satisfy the three psychological needs, it cannot be concluded that the interactive e-book was more effective at cultivating need-satisfaction than the traditional instruction and flipped classroom models implemented in this study because need satisfaction is not controlled for.

Zhao et al. (2021) collected results through tests, questionnaires, and student interviews. First, all participants were given a pretest with the purpose of evaluating whether or not students in each group had similar prior knowledge of whole numbers. The pretest consisted of 25 questions and was developed by three math teachers. The pretest results were not reported in this study. Then, students were given a pretest questionnaire so that the research team could track the motivation and metacognition levels in all participants. With the foundation laid, each group moved into the core activity of the study which lasted a week (90 minutes of total participation). The same teacher was used to facilitate each group's learning process (Zhao et al., 2021).

Zhao et al. (2021) made sure the experimental group and the control group learning in a flipped classroom setting had a similar experience interacting with the material, with the only fundamental difference being that the experimental group solved problems in the interactive e-book. Both groups had group discussions, quizzes, and presented their learning at the end. If students were confused about the content, the teacher was there to answer questions and guide the learning process. In contrast, the traditional classroom group read in their textbook before they were lectured on fractions (Zhao et al., 2021). Clearly, students in the experimental group and the flipped classroom group had a more interactive experience. Conventional wisdom suggests that a more interactive experience leads to an increase in achievement and motivation, but what do the results suggest? Does the mystery of a snowy castle distract students from learning fraction basics?

After each group took a posttest 50-minute, 20-question posttest, Zhao et al. (2021) calculated the mean score of all three groups. Then they calculated the adjusted mean score to account for outliers in data without arbitrarily removing the extreme data points. The interactive e-book group's adjusted average is 65.83, while the flipped and traditional group's adjusted average is 61.84 and 41.49, respectively. After further statistical analysis, the adjusted mean score of the e-book group was significantly higher than the adjusted mean of the other two groups (Zhao et al., 2021). In other words, there was a low chance the adjusted means of all the groups are equal. Therefore, the adjusted means are likely different. Notice this was not the same as concluding that the arithmetic difference between the means is significant.

In order to answer the remaining questions, namely whether or not motivation and metacognition were at the highest levels in a gamified classroom, Zhao et al. (2021) analyzed the results from a posttest questionnaire that consisted of five items assessing metacognitive tendency and five items assessing learning motivation. The results indicated that the students in the gamified and flipped classroom had a higher adjusted mean metacognitive score than students learning traditionally, and this difference is statistically significant. Moreover, the adjusted mean scores of the gamified and flipped classrooms were within six one-hundredths of a point of each other, and there was no significant difference. Similarly, the results of the learning motivation post-questionnaire show that students learning with the interactive e-book were more motivated than the other two groups (Zhao et al., 2021).

Finally, Zhao et al. (2021) interviewed 12 students (four from each group) to get their perspective on their experience as a whole and its subsequent impact on learning and motivation. They scored 16 different key phrases from student responses that relate to performance, motivation, self-reflection, and acceptance. Of these, only three instances of students in the flipped classroom group scoring higher than students in the gamified group occurred. The traditional learning group's only score was in the self-reflection category. This points to a substantial difference in the learning process, according to students' perceptions of their own learning. In general, students in the gamified and flipped classroom groups agreed that the pre-learning activity, namely watching a 15-minute video on fractions, was helpful to their in-class learning. Students in the gamified group liked being able to go back to the interactive e-book for clarity when they encountered a problem they did not understand (Zhao et al., 2021). The differences in the mean posttest scores may partially be explained by this added layer of flexibility.

In summary, Zhao et al. (2021) wanted to know what impact combining a flipped and gamified classroom has on student achievement, motivation, and metacognition as it relates to learning fractions. Their research boiled down to three fundamental metrics: the posttest, post-questionnaire, and interview. To provide context for the conclusions, other students learned fractions in a conventional flipped and traditional learning setting. The results of posttest scores showed that mean student achievement and motivation were the greatest in the blended learning group, followed by flipped learning and traditional classroom. Moreover, the interview results showed that, in general, students in the blended and flipped group *perceived* themselves to be more motivated than the traditional learning group Zhao et al. (2021).

Virtual Reality

Shi et al. (2022) created a game-based virtual reality learning environment (GIVRLE) that facilitated the learning of quadratic functions. Unlike other gamified models, students were able to walk around in the game while they wore the virtual reality headset. Through the headset, students were able to play *The Crazy Pot* game, which required them to understand the properties of quadratic functions to break pots with projectile launchers. The basic idea is that the path of the projectile models parabolic motion. Students must use the location of the pots and an incomplete quadratic equation to break the pots with rocks. The quadratic equations become increasingly difficult as students moved through the four levels. The GIVRLE has a built-in feedback system such that students answering a problem incorrectly three times in a row triggers a helpful fairy that subsequently flies over to an animation screen that walks students through the general process. The walkthrough process is scaffolded in that the walkthroughs give more clues as mistakes increase. Finally, if the students passed all levels they get to shoot balloons with a bow and arrow while riding a bird (Shi et al., 2022).

In total, 103 seventh graders from two different schools in China were selected by Shi et al. (2022) to participate in this study. The students were separated into two different groups such that the students in the experimental group played the game in the GIVRLE and the control group learned quadratics as they would have in a traditional classroom setting. First, both groups were given a pretest and pre-motivation survey, which consisted of five and four questions, respectively. Students were scored based on the number of correct answers for the pretest and posttest, whereas motivation was scored based on students' perceptions of the importance and difficulty of math on a Likert scale. Next, all students learned quadratic equations in their respective groups. The experimental group played the game for 20 minutes and the control group spent 40 minutes taking the posttest after learning the material (Shi et al., 2022).

Shi et al. (2022) evaluated the possibility that GIVRLE may be a great tool for increasing student achievement and motivation by comparing mean scores of the pretest, posttest, and learning motivation surveys within each group to track progress. The results showed that the experimental group's mean score on the posttest was about 12% higher than the mean score on the pretest. On the other hand, the control group only had a 4% improvement, though they scored about 21% higher on the pretest than the experimental group. So, the research team concluded that playing the game was helpful to quadratic function learning. Moreover, the learning motivation scores of the experimental group showed a 2.23-point increase in the posttest (Shi et al., 2022).

This answers the guiding research question of this thesis; gamified learning does have a positive impact on student learning and achievement. Beyond the scope of any study, but surely begging the question is this impact enough to justify the cost of virtual reality headsets? The point of reporting these studies in detail is to help educators build such justifications. Then it is fair to point out some of the study's limitations. The study conducted by Shi et al. (2022) was small because of the nature of the gamified medium; the students in the experimental group played the game one at a time. The size of the study may have had a snowball effect on interpreting the improvement between posttest scores. For example, another class selected to participate in this study could have had higher pretest scores. The small sample size does not guarantee that every class will have a 12% increase (Shi et al., 2022). Unfortunately, it was impossible for them to have the students in the control group, which had significantly higher pretest scores, play the game and retake the posttest. So, extrapolating these results to other classrooms is difficult but, on the other hand, there is no evidence to suggest that student achievement and motivation would be negatively impacted.

Touch-Screen Intractability

The prior studies discussed primarily relate to games that use visuals to supplement the narratives, creating a more immersive experience; however, one of the undisputed strengths of gamification in math classrooms is that it allows students to visualize abstract concepts. This is precisely what Ottmar et al. (2015) had in mind when they designed *From Here to There!* (FH2T), an interactive environment designed to enhance the understanding of middle school students learning the properties of algebra. Students exploring FH2T find themselves immersed in a "universe" containing 14, "worlds", each covering a unique topic and containing 15-20 questions. Students must answer most questions correctly to traverse to the next world. The novelty is seen through the intractability as students can move numbers around with their fingers as they use the properties of algebra to match the starting expression with the desired expression (Ottmar et al., 2015).

Before conducting an experiment, Ottmar et al. (2015) created a pilot study in which 110 middle school students (six classes) engaged in FH2T for a total of four hours across six class periods. Students encountered problems in two modes, namely "fluid" and "retrieval." In retrieval mode, students had to enter the correct answer for algebraic and arithmetic transformations, whereas in fluid mode the software calculates the correct answer when students select the appropriate operation. Upon gathering posttest data, there was an 8.5% increase in mean score from the pretest. Given the absence of an experimental design definite conclusion cannot be made. There is no control group; normal classroom learning, intermixed with FH2T sessions, could have accounted for the increase in posttest scores (Ottmar et al., 2015). Nevertheless, the results indicated that FH2T was trending in the right direction. Consequently, they proceeded to discover the capabilities of FH2T further by conducting an actual experiment, which controlled for increases in achievement from normal classroom instruction.

Ottmar et al. (2015) assigned five classes of middle school students (N=85) to three groups, namely retrieval practice, fluid visualization, and the regular instruction control group. Notice the division of the experimental group into two subgroups. So, in addition to testing the efficacy of FH2T, they wanted to test different features of FH2T. The students took a 30-question pretest before participating in the experiment, which spanned six class periods and lasted about six hours. The experimental groups had an identical experience with respect to the worlds they explored and the questions they were asked. The fundamental difference between these two groups was how they must solve each problem. The fluid visualization group needs only to identify the correct operations for each step of the problem because the software computed each operation for them. Alternatively, the retrieval group had to identify the correct operations, move symbols around accordingly, and input their answers (Ottmar et al., 2015).

Since the elements and mechanics of FH2T were documented thoroughly, looking for the conditions of need satisfaction is possible. First, autonomy can be satisfied if students want to play the game, but it should be noted that there are no major, in-game choices to be made. Autonomy fails to be satisfied for students that would not choose to play FH2T on their own. Second, competence can be satisfied through students earning the right to solve problems in fluid visualization mode by completing some of them in retrieval mode first. Moreover, students may feel accomplished by earning points and unlocking new worlds to explore. Finally, relatedness is not apparently satisfied, because students are working individually. Of course, this need can be met through other means, so it is important to use gamification as a supplementary resource. Again, the lack of a need(s) being satisfied (autonomy and relatedness in this case) does not disqualify FH2T from being a motivator; however, we would have no preliminary reason to predict that motivation will increase.

The focus of this study was narrowed to two main research questions. First, do students playing FH2T have a greater improvement in posttest scores than those learning normally? Second, does more exposure to FH2T predict improved performance and learning? Put another way, is a student's progress in the game reflected in their posttest score? The first question was answered after an analysis of covariation to predict posttest scores as a function of their group, with gender and pretest scores as covariates. Ottmar et al. (2015) found the differences between the groups to be significant. Hence, they adjusted the mean score of posttest results for all the groups. The results showed that students scored higher in the fluid visualization group than the other groups, and the difference is statistically significant. There is no significant difference in mean score gains between the retrieval and the control group, in fact, the mean posttest scores were lower.

The second question was answered by the data collected from the experimental groups. Ottmar et al. (2015) found that, on average, students in the retrieval group completed about four “worlds”, whereas students in the fluid visualization group completed about six “worlds”. Then they conducted a hierarchical linear regression model which indicated that for every world a student completes their posttest score increased by .76. Finally, do high pretest scores plateau posttest scores and marginalize the efficacy of exposure? In short, regression analysis showed that the answer is yes. The data was grouped by pretest score (low, mid, high) and showed that within each grouping, more exposure time led to a higher posttest score; however, more exposure did not always compensate for a lower pretest score across groups. This translated to many mid-and high-group students scoring higher on the posttest than anyone in the low-pretest group. So, while there was a .76 increase in posttest score per world completed, this was not enough for some students to overtake their peers with much higher pretest scores (Ottmar et al., 2015).

In summary, Ottmar et al. (2015) found that the fluid visualization group had a substantial gain in posttest score of about one-third of a standard deviation; the other two groups saw small, decreased gains. Why the retrieval group did not have similar results to the fluid visualization group is unclear. It is possible the more interactive and engaging format of the fluid visualization group was responsible for the better results. Still, maybe it was the fact that the retrieval group had less exposure because their calculations took extra time. Then, it is possible that the retrieval group never learned some concepts that appeared on the posttest (Ottmar et al., 2015). Nevertheless, these results are not entirely as negative as it appears for the game as a whole. Recall that the two experimental groups were derived from another version of FH2T which allowed students to earn fluid visualization after completing a fixed number of problems by retrieval method. The overarching goal of this experiment was to see which mechanisms of FH2T should be prioritized moving forward. This study is important because it answers the student achievement portion of the research question in the affirmative and highlights the importance of gamified learning actually providing an interactive environment. Four years later they conducted another study, identifying other important game mechanisms.

This time, Hulse et al. (2019) adapted FH2T to fit an elementary school audience, resulting in the new game, *From Here To There: Elementary* (FH2T:E). For the purposes of the study, they made another version of FH2T:E without gamified elements, such as points, awards, and recognition of progression. Their inquiry boiled down to three main questions. Primarily, they wanted to know if there is a difference in learning in the gamified and non-gamified versions of FH2T:E. Second, can in-app logs, monitoring the problem-solving process, predict learning outcomes? If so, this would be a great feature to implement to help teachers support in-game learning. Moreover, do certain behaviors in FH2T:E predicts learning for high and low students? To answer such questions, they selected 185 students, from an elementary school in Massachusetts, to participate in the study by playing the game for four, 20-minute intervals. The students were further split up into two groups; one played the gamified version and the other played the non-gamified version (Hulse et al., 2019).

In regard to the first research question Hulse et al. (2019) said, “Results suggest that there were no differences in posttest performance, between the gamified and non-gamified condition, when only condition, gender, and pretest performance were used to predict posttest performance.” (p.434). In other words, no inferences could be drawn from any given posttest score as to whether they were in the gamified or non-gamified group, were male or female, or performed well on the pretest. After analyzing in-app log data, however, on average, students in the gamified group scored 6.58 points higher on the posttest than students in the non-gamified group. But, the only predicting factor for posttest performance was progress. Students who completed more problems in the app scored higher on the posttest. For each standard deviation increase in completion, there was approximately a 3.07-point increase in posttest score. Finally, students who completed more problems, but had low pretest scores, had more learning gains than students who completed fewer problems. Similarly, lower students who engaged with more problems had more learning gains. Engaged with more problems, in this context, refers to students who attempted problems multiple times, either because they were incorrect the first time or went back for more practice and the amount of time spent interacting with problems. Yet, the number of problems completed and engagement with problems did not affect academic performance for students who scored high on the pretest (Hulse et al., 2019). This seems intuitive as high performance on the pretest is indicative of already knowing most of the content. Also, high-knowledge students may be more intrinsically motivated to learn already, and thus non-math gamified elements would be redundant or only have a marginal effect.

Hulse et al. (2019) boast accessibility as one of the strengths of FH2T:E. The results from monitoring the in-app progression of low- and high-knowledge students showed that both groups easily progressed through the game. Students were able to play FH2T:E relatively anxiety-free, because there was no internal timer. One of the main takeaways from their research is that predicting factors, which correspond to different mechanisms in games, can be correlated with student achievement. Overall, five factors were found to predict student achievement: engagement, progress, strategic flexibility, strategic efficiency, and speed. But some factors may not be correlated, so developers must continue to refine their games and test these intricacies (Hulse et al., 2019). Teachers serve to gain much if games allow them to monitor such factors, and clearly, this is not an option in a non-gamified classroom. Then, there is much-needed research to be conducted pertaining to the innermost workings of games, because most studies paint with a broad brush, they may miss the effects of important in-game mechanisms.

Adaptive, Competitive, and Collaborative Gamification

In order to have a more robust appreciation of gamified learning, it helps to have a broader understanding of the many different varieties. Jagust et al. (2018) wanted to learn more about competitive, adaptive, and collaborative gamified learning and contrast them with a traditional learning environment. The research of Jagust et al. (2018) boiled down to two fundamental questions. First, how do different types of gamification affect student performance in primary (elementary) school? Second, how do the defining game elements in different types of gamification add to student performance? Two second-grade and one third-grade class were selected for the purpose of this study from a Croatian elementary school. The study was segmented into four parts, such that each participating class experienced a circuit encompassing traditional learning and each kind of gamified learning. Each segment lasted approximately 15 minutes and no two were scheduled for the same day. Presumably, they were still learning between the scheduled segments as normal, which is not to be confused with traditional learning. Additionally, each of the segments catered to a topic they would normally be learning, namely adding and subtracting numbers from zero to 100 and multiplying and dividing by numbers two through five and zero (Jagust et al., 2018).

Jagust et al. (2018) made the learning experience of each segment was similar in that all students participated through a math widget that was modified uniquely for each experience. Notably, the students who participated in the traditional segment still answered questions through the widget. The data was collected by the widget to calculate the average number of correct attempts, incorrect attempts, and total time taken to complete all questions. Unlike previous studies, students were not given a pre and posttest to base conclusions about student achievement on. Thus, answering the guiding research question pertaining to student achievement will require a deeper dive into the specifics of each widget modification. It is beyond the scope of the research conducted by Jagust et al. (2018) to mention teacher-usability factors, but presumably, adjusting the settings of the math widget to accommodate a chosen variety of gamified learning is not difficult.

Though Jagust et al. (2018) did not explicitly talk about need satisfaction, need satisfaction occurred, to varying degrees, in all three varieties of gamified learning. Relatedness is mostly satisfied in the adaptive and competitive varieties of gamified learning when students are working together to defeat the computer virus and compete against their classmates for the best score. Competence is arguably satisfied in all three varieties, especially in the adaptive version because the winning conditions are adjusted based on the performance of each student. The need for autonomy can also be satisfied in all three varieties if students are willing to play the game. Given the interesting nature of the virus scenario, this is a likely possibility. All things considered, the presence of need satisfaction corroborates the conclusion that student achievement and motivation is increased as a result of gamification.

Modifications to the math widget define distinct learning experiences through gamification, though commonalities exist throughout. Jagust et al. (2018) had all participants answer mathematics questions on the widget, were required to retry wrong answers, and participated no more than 15 minutes. Basically, students participating in the non-gamified variety participated at the base level of the widget: answering arithmetic questions without a timer (Jagust et al., 2018). The competitive gamified learning was facilitated by projecting a leaderboard in front of the class, which updated continuously, displayed a 15-minute timer, and showed a running total of the top six participants (Jagust et al., 2018). Displaying a public leaderboard uniquely defined the competitive gamified learning experience.

Jagust et al. (2018) made the adaptive segment with two important elements in mind, namely a narrative and a built-in algorithm that personalizes the experience. The design of the narrative was meant to be engaging. Students must answer questions correctly, before an evil computer virus does, to protect the school webspace from catastrophe. Embedded in the game is an algorithm that is designed to give students a continuous challenge based on their performance. If students are consistently beating the virus in answering questions, their time on subsequent questions is reduced by one second. On the other hand, if students are lagging behind the virus' progress, students are given more time to answer questions before the virus does. In case such adaptations were insufficient in ensuring victory, during the last three minutes of the game a reasonable amount of time was added on the students' side, more than enough time to defeat the virus (Jagust et al., 2018).

The collaborative segment combined elements of the other varieties of gamified learning, in that there was a projected leaderboard and narrative elements (Jagust et al., 2018). This time, the sum of student points was totaled against the sum of virus points and projected on the leaderboard (Jagust et al., 2018). To make sure students gave their best effort throughout, an illusion of urgency was created by having the virus' point total set at about 90% of the student total (Jagust et al., 2018). Clearly, students who perform better under less pressurized conditions will enjoy collaborative gamification.

After all three classes completed the circuit of gamification, Jagust et al. (2018) applied statistical analysis to base their conclusions. Jagust et al. (2018) found that all three varieties of gamified learning led to improved performance, or at least, plateaued results. Performance was measured with respect to time and the number of correct answers. "On the other hand, the activities in the non-gamified condition resulted in negative performance levels, with students becoming bored and/or starting to lose focus as the noise levels and jitter in the classroom increased" (p. 454). While there were some performance dips during the non-gamified segment, taken as a whole, the differences in performance between competitive, collaborative, and non-gamified learning were not found to be statistically significant (Jagust et al., 2018). This may suggest more defining characteristics than just a leaderboard and timer ought to be added to the competitive experience; however, competitive gamified learning may just be too polarizing for the entire class to benefit from. It is somewhat surprising that the collaborative version did not yield significant differences, but the results among classes in this category varied markedly. Adaptive learning yielded the highest performance and its scores were significantly different than the non-gamified version. The built-in, personalized algorithm helped students, who might have otherwise quit when the virus was winning, stay in the game, and students, who would have otherwise been winning by a large margin, were able to stay engaged and answer the rest of their questions with effort (Jagust et al., 2018).

Though the results were encouraging, Jagust et al. (2018) was honest about the negatives inherent in conducting experiments like this. First, the sample size (N=54) is too small to generalize results to other classrooms. This was evident even within the experiment because the standard deviation in the mean number of correct answers was no less than 30 points across all categories. Whether or not these games sustain engagement for long periods of time is left unanswered; studies that require the participation of longer than 15 minutes could be used to explore this capacity. Moreover, the study uses a unique combination of gamification elements that make replicating it in another game a challenge. Finally, no longitudinal efforts or follow-up is evident Jagust et al. (2018). Longitudinal studies might provide some answers to important questions pertaining to the appropriate amount of gamification time allotted in a given unit, the number of different games that ought to be used to sustain engagement and broader ramifications of gamifications' impact on student achievement.

Rewards-Based Gamification

Non-academic game elements, like scoring, timer, and leaderboard, are commonplace in gamified learning, but they could actually be counterproductive in eliciting motivation. Elements also may have different effects on different learning styles. Reyssier et al. (2022) categorized 258 students (twelve classes) from four high schools into six player profiles, defined by the HEXAD gamification user typology, based on their answers to the HEXAD questionnaire. Participants also took a pre-motivation questionnaire called the Academic Motivation Scale, which yields subscores for seven different kinds of motivation. After learning the algebra lesson and playing the game, Reyssier et al. (2022) compared pre and post-test motivation scores using a Wilcoxon signed ranked test, which yielded a significant decrease in intrinsic motivation and external regulation and a significant increase in the amotivation of students. Then, Reyssier et al. (2022) used partial least squares path modeling to see what effect initial motivation scores and player profile have on the changes in motivation. They concluded that “34.3% of the variation in intrinsic motivation, 12% of the variation in extrinsic motivation, and 45.1% of the variation in amotivation, could be accounted for by the level of initial motivations and the learner player profile” (p.7). So, extrinsically motivated students, who are most likely to be affected by game elements, may need these elements to engage at a high level with the material. The opposite is true for intrinsically motivated students. Moreover, there were significant differences in motivation and achievement depending on what game element students were randomly assigned. For example, students who were assigned badges had a higher ratio of correct answers than students who were given a timer. Students with a timer reattempted fewer quizzes than those with other game elements (Reyssier et al., 2022). Thus, students with other game elements were particularly motivated to improve their quiz scores.

One of the most important implications of this study stems from the fact that game elements can be recommended to students based on their learner profile and initial motivation results. A makeshift zodiac for the different player profiles emerges. Reyssier et al. (2022) recommends badges for disruptors and students who are intrinsically motivated. Progress is helpful in building intrinsic and extrinsic motivation, so this game element can be recommended for amotivated learners. Ranking is helpful to students characterized as free spirited, but tends to demotivate highly motivated and amotivated students. Score is recommended for socializers, but disruptor philanthropists are likely to me demotivated. Generally, score is not recommended for motivated or amotivated learners. A timer is recommended for achiever and free spirit profiles;

however, philanthropists and intrinsically motivated learners will find this demotivating. Avatar is recommended for amotivated students, but socializers may find avatar demotivating. In conclusion, how students are motivated matters (Reyssier et al., 2022). Educators ought to look for a specific subset of game elements that keep amotivated, intrinsically, and extrinsically motivated students engaged during the lesson and modify games accordingly.

Rincon-Flores et al. (2023) developed a user-friendly, web-based platform called *Gamit!* to help teachers implement gamified learning. Through *Gamit!*, teachers awarded badges, managed leaderboards, and reviewed a final performance report, and students selected an avatar and saw themselves on a leaderboard. Rincon-Flores et al. (2023) wondered what effect rewards-based gamification has on the attitudes and attention, engagement, and resilience dimensions of high school math students. Anxiety, enjoyment, usefulness, motivation confidence, and procrastination were selected as attitude dimensions. Four hundred and fifty-four (454) Mexican high school students were selected to participate in this study, in which *Gamit!* was integrated into a semester-long course. In order to answer the question about attitude, students were given a pre- and post-questionnaire scored on a Likert scale. Similarly, students were given another questionnaire at the end of the course which was then scored according to attention, engagement, and resilience dimensions (Rincon-Flores et al., 2023).

To answer the attitude question Rincon-Flores et al. (2023) used a questionnaire that yielded scores in the six selected resilience dimensions. The results showed that management of anxiety, enjoyment, and usefulness increased, whereas motivation, confidence, and management of procrastination decreased. The results of the second questionnaire indicated that students viewed themselves as more attentive, engaged, and resilient by the end of the course. Moreover, students studied at home more; sixty percent of students studied more than two hours at home, compared to 48% of students before the course (Rincon-Flores et al., 2023).

CHAPTER III: CONCLUSION

Summary of Literature

This thesis reviewed nine studies in depth to answer the guiding research question: how does gamified learning increase motivation and achievement in the math classroom? The research teams that authored the reviewed studies created unique games for students to play during the experiment. Most of the studies lasted days to weeks. Only Rincon-Flores et al. (2023) conducted a study that lasted the entire duration of a normal course. All studies in Chapter Two explored gamified learning's effect on motivation, except Li et al. (2022), in varying depths. Studies that involved motivation based the conclusions on different data collection methods like interviews and self-report surveys taken by students (Jagust et al., 2018; Reyssier et al., 2020; Rincon-Flores et al., 2023; Shi et al., 2022; Zhao et al., 2021) or inferences on engagement levels based on in-app data (Hulse et al., 2019; Ottmar et al., 2015). Additionally, Rincon-Flores et al. (2023) asked teachers to report on their observations about motivation.

Similarly, all studies explored gamified learning's effect on student achievement, except Rincon-Flores et al. (2023). Student achievement results were primarily determined by analyzing the change in pre-and posttest scores.

Motivation Results

Does gamified learning increase motivation? The research studies reviewed in this thesis report the findings in two main ways, namely, self-report questionnaires and application data. Self-reports yielded more straightforward conclusions about the effect of gamified learning on motivation, but analyzing in-app data is a less biased approach. Unfortunately, the research teams that collected data from the application were concerned mostly with student achievement, and consequently, shared the results through that lens. Nevertheless, there were clues in the application data that showed some students were very engaged in the games, which suggests playing games increased their motivation levels. The results based on objective, in-app data reports will be discussed first, followed by the self-report findings.

In-App Data reports

Jagust et al. (2018) used in-app data to make inferences about the engagement levels of students and found that the groups using the adaptive algorithm, which adjusted difficulty according to accuracy, were the most engaged. Similarly, inferences about motivation can be made from the in-app data from the game *From Here to There!* created by Hulse et al. (2019) to analyze student motivation. Regression analysis illustrated that as time spent in the application increased, students in the non-gamified group solved less and less problems. The gamified groups all solved more problems as in-game time increased. Students in the gamified group were engaged throughout and their motivation levels had probably been maintained or increased, in spite of mental fatigue, to solve more problems than the non-gamified group as time passed. On the other hand, diminishing motivation levels could explain the dip in performance of the non-gamified group (Hulse et al., 2019).

Self-Report Surveys

There were three studies using self-report surveys that concluded unequivocally that the gamified learning group had gains in motivation greater than the traditional instruction group (Rebollo et al., 2022; Shi et al., 2022; Zhao et al., 2021). For example, Rebollo et al. (2022) found that 67% of students would be willing to spend more time doing this activity. Reyssier et al. (2020) and Rincon-Flores et al. (2023) had mixed results.

Reyssier et al. (2020) broke down motivation into three categories: intrinsic motivation, extrinsic motivation, and amotivation. Upon comparison of the results of the Academic Motivation Scale questionnaires, which were given to students before and after the experiment, it was found that students who were initially amotivated were more intrinsically motivated by the end. Conversely, students who were initially intrinsically motivated were less intrinsically motivated by the end. The same was true for students who were initially extrinsically motivated (Reyssier et al., 2020). Furthermore, the motivation results varied for different learner profiles depending on the element they were assigned. After correlating learner profiles with motivation types, Reyssier et al. (2020) found that the achiever was significantly correlated with all types of motivation, whereas the philanthropist had the weakest relationship to all types of motivation.

After Rincon-Flores et al. (2023) compared the results of the pre-and-post questionnaire, it was found that students experienced negative gains in the motivation category of the questionnaire. Yet students studied longer and the professors noticed motivation improve. The professors were interviewed and said that it was obvious motivation improved over the span of the course, as evidenced by the participation in class (Rincon-Flores et al., 2023). The mixed results could stem from students not feeling motivated but still participating in class, which may suggest playing the game only increased extrinsic motivation.

Achievement Results

Does gamified learning increase achievement? Before answering the second part of the guiding research question, it is important to provide the context of the results as defined by the changing experimental conditions. The research studies highlighted in Chapter Two did not all control for the same condition, which makes a broader summary misleading. Essentially, the research teams either controlled for traditional learning, using the same app without gamified elements, or had a valid reason not to include a control group.

Controlled for Non-Gamified Learning

Hulse et al. (2019) did not control for traditional learning, rather a version of From Here to There (FH2T) was created without game-like features (e.g., level progression). No statistically significant difference in posttest scores was observed between the two conditions. After the in-app data was analyzed, however, students who progressed faster through the app tended to have higher posttest scores, and this difference was close to being significant (p=.056). Digging deeper, of the students who were in the lowest grouping of pretest scores, the students who had completed more problems had the higher posttest scores than those with less completed problems. This stands in contrast to students in the highest grouping of pretest scores who had similar posttest scores regardless of how many problems were completed. Similarly, the low pretest score group benefited from going back and solving problems a second time, whereas the high pretest score group did not see this effort pay off in the posttest (Hulse et al., 2019). Since there was no statistically significant difference in posttest scores between the two conditions (game-like and plain) when in-app data was not taken into account, it follows that the game-like features of FH2T are irrelevant to achievement. Given that these features are standard features in many educational games this is a relevant finding. This is consistent with the findings from the application data: student achievement was highly influenced by the prior knowledge students had, as evidenced in the pretest scores (Hulse et al., 2019). The high-achieving students did not gain anything from the game elements but low-achieving students' posttest scores were correlated with number of problems attempted and number of second attempts (Hulse et al., 2019). Since problems could be attempted or reattempted for each version of the application, the game-like features were probably irrelevant. These conclusions seem to contradict the fact that posttest scores of the gamified group were on average six points higher than the plain-app group. This could be explained by the limitations of quasi-experimental studies, and more specifically the lack of mirroring between the groups.

As discussed earlier in Chapter Two, Jagust et al. (2018) conducted a study with three different groups that each used a game with a different gamified dynamic: adaptive, collaborative, and competitive. The dynamics were pseudo-controlled against a version of the software stripped of game elements. Instead of having one control group, participants cycled through each dynamic. The adaptive game was defined by its capability to challenge students by adjusting the total time remaining to beat the virus (solving problems accurately) according to how accurately they were solving problems, with more time being added for the struggling students and time being subtracted for the thriving students. Students answered more questions correctly and incorrectly in the adaptive group (Jagust et al., 2018). In this case, student achievement increased more in the adaptive group, because the adaptive game dynamic contained more meaningful qualities than the rest of the groups, which either had superficial elements, such as narratives and leaderboards, or no gamified elements.

Controlled for Traditional Instruction

Four of the research teams controlled for traditional instruction (Ottmar et al., 2015; Rebollo et al., 2022; Shi et al., 2022; Zhao et al., 2021). Three of the research teams shared results which showed that the gamified group outperformed the traditional instruction group (Ottmar et al., 2015; Shi et al., 2022; Zhao et al., 2021). Rebollo et al. (2022) found no statistically significant differences in learning.

Professional Application

There are two choices educators can make after reading this thesis. In the case that the information is found satisfactory and educators want to implement gamified learning in the classroom, careful planning should be done to ensure an optimal experience. First, common elements and mechanics that appear in many games should be understood so educators can decide which combination of elements best suits the class. Second, educators should find games that provide extensive summary statistics. Finally, conclusions from this thesis about superficial elements, like badges, leaderboards, and avatars, which may have varying degrees of utility for different ages, should be understood as primarily coming from elementary and middle school demographics. Elements may have great effect in one age group yet become irrelevant as the child ages. Mechanics like touch-screen intractability, adaptive algorithms, or any visualizations features seem to remain relevant as the students age.

In the case that educators do not feel comfortable generalizing the conclusions in this thesis to their class, important takeaways still exist. First, SDT is a valid theory with many applications beyond gamified learning. Second, the conclusions on motivation and achievement could be interpreted as mixed; however, it is a mixture of neutrality and positive gains. None of the studies reviewed suggested that gamified learning erodes the will power or achievement of students irreparably. Thus, ceteris paribus, there is little to no downside to trying out gamified learning. The opportunity cost of having students play games may be steep, but this can be minimized, for example, by only using games when the class would normally be on a break.

Limitations of the Research

This literature review consists of 30 references and is not an exhaustive review of all literature on gamified learning in math classrooms. The limited number of references brings concerns about the generalizability and accuracy of the results and, subsequently, the conclusions. In this case, educators may want to generalize results from gamified studies to their own classrooms and need the results to be accurate in order to justify the expense of implementing expensive technology to the administration. Self-reports, quasi-experimental studies, and applying SDT all have inherent limitations that will affect generalization and accuracy of results.

Self-Reports

Self-report bias occurs when people answering questionnaires make misrepresentations to appear more favorable or answer inaccurately because they do not know the answer (American Psychological Association, n.d.). Yet conventional wisdom suggests that it is reasonable to go directly to the source to get information, and according to McDonald (2008) that is one reason why self-reporting methods are so widely used to collect data. Additionally, they are a fast and relatively inexpensive way to collect data (Kline, 1993 in McDonald, 2008). However, there are a number of reasons why self-report questionnaires can be unreliable. Schwarz (1999) argued that self-reports are fallible sources of information because minor changes to the wording, format, or context of a question can result in significant changes in response. The remedy to this is to use a variety of quantitative and qualitative methods (McDonald, 2008). Some studies reviewed in the previous chapter used self-report questionnaires to track gains in motivation (Rebollo et al., 2022; Reyssier et al., 2020; Rincon-Flores et al., 2023; Shi et al., 2022; Zhao et al., 2021), but none of these studies used quantitative methods to draw conclusions about motivation. On the other hand, Hulse et al. (2019) used game data to analyze factors that influence motivation, such as progression and engagement but did not use self-reports.

More specifically, students may feel pressure to answer more favorably on surveys as they attempt to align themselves with positive cultural norms, a phenomenon known as social desirability (Krumpal, 2013). Furthermore, teachers involved in research studies may be more motivated than average to provide greater motivational support to their class (Flunger et al., 2022).

Quasi-Experimental Studies

A quasi-experimental study is a like a normal experimental study in that it is a type of quantitative research that involves changing the independent variable and documenting the subsequent effect on the dependent variable; however, this type of study does not randomly assign participants to groups and may not have a control group (Aloe et al., 2017). Applying this definition to the studies reviewed in this thesis, it follows that most were quasi-experimental on the basis that entire classes were assigned to either the control group or experimental group. According to Scher et al. (2015) researchers who conduct actual experiments can be more confident that the results of their respective studies are only measuring the effects of the independent variable. With that being said, quasi-experimental research studies are widespread in education and are certainly not to be dismissed as useless. Moreover, mixing classes and taking students out of their normal learning environment could mean adding confounding variables to the experiment anyway.

Not all quasi-experimental research studies are equal; there are a few precautions researchers can take to make sure the results are more accurate. First, Scher et al. (2015) says researchers ought to find groups that mirror each other; the closer the experimental and control group are to being identical in gender, race, ability, and prior knowledge, the more accurate the results are going to be (Scher et al., 2015). Li et al. (2022) used this principle when they selected an equal number of males and females from two high-achieving classes in an urban middle school in Shandong, China. The rest of the studies had roughly equal gender participation as well.

Second, Scher et al. (2015) said the experimental and control groups may differ by unseen variables like motivation. Therefore, selecting groups from similar demographics and settings is ideal. For the most part, the research teams highlighted in this thesis followed this principle; however, the students enrolled in the classes selected to participate, for example, were not necessarily members based on equal achievement or motivation levels. Conclusions based on motivation are probably not as biased in studies that controlled for baseline motivation levels by giving a pretest questionnaire; however, there are many variables that affect student achievement that are not necessarily mirrored in the other classes participating in the study. Thus, aggregate test scores and subsequent conclusions may not be accurate.

Small Sample Size

When designing a research study, calculating the appropriate sample size is necessary in order to get statistically significant results (Burmeister & Aitken, 2012). If researchers made calculations, none were included in the journal articles. Many research teams appeared to use the minimum number of classes required to have a functioning experiment, to the order of every grouping only containing one class (Jagust et al., 2018; Li et al., 2022; ; Shi et al., 2022; Zhao et al., 2021; Rebollo et al., 2022). For example, Rebollo et al. (2022) selected two classes (37 students) to participate in their study, enough for one class in the experimental and control group each. Some research teams were presumably concerned with sample size; they had relatively large sample sizes that incorporated multiple classes (Hulse et al., 2019; Ottmar et al., 2015; Reyssier et al., 2020; Rincon-Flores et al., 2023). Notably, Reyssier et al. (2020) used 12 classes from four different schools, and Hulse et al. (2019) used 10 classes from three different schools.

Limitations of Applying SDT

Determining whether the psychological needs of participants in research studies are being satisfied can be difficult. Generally, studies that offer limited descriptions of the proposed game do not allow for any analysis of need satisfaction. Even if the elements and mechanics of the game are properly described, SDT analysis cannot yield quantitative results if need satisfaction is not explicitly addressed in self-reports; the number of students who had autonomy, competence, and relatedness satisfied is unknown. The most that can be concluded is that all needs are able to be satisfied by playing the game. Moreover, if motivation increases after implementing some game, it is beyond the scope of most of the studies reviewed here to interpret the results in terms of intrinsic and extrinsic motivation, which can yield more mixed results.

Many studies have validated SDT by showing that intrinsic motivation and autonomous extrinsic motivation impact student engagement and academic achievement positively (Niemiec & Ryan, 2009) Even so, there is an absence of long-term studies in this area; many studies are cross-sectional and do not control for baseline achievement (Taylor et al., 2014). Few studies follow up with their participants over time and do not measure changes in achievement that may be caused by motivation (Taylor et al., 2014).

Moreover, the conditions surrounding the experimental design, where data has been collected from student performance on designed assessments or self-report questionnaires, may artificially satisfy needs.

Implications for Future Research

The short answer to the guiding research question is: gamified learning can increase motivation and achievement given the implementation of a game with a meaningful combination of elements and mechanics, but it is not clear what impact any given game-like element has. According to Jagust et al. (2018), one of the biggest challenges facing gamified learning is finding the right combination of elements and dynamics to engage students for a long period of time. Two distinct implications for future research are represented in this statement. First, more research needs to be done to explore meaningful elements and mechanics, like adaptive algorithms, touch-screen intractability, nuanced narratives, and to discover new ones. Second, studies should have students playing the game for longer periods of time in one sitting. Furthermore, researchers should seek to isolate elements like Resssier et al. (2020) did to ascertain the individual effect on motivation and achievement. Then educators and game designers would know which elements are foundational to gamified learning.

More studies ought to be implemented with a longer-time frame, especially studies that monitor motivation and achievement spanning the entire course. Course-long studies would allow researchers to pull data from actual tests, activities, and final grades. Nevertheless, if their own posttest was used, the results could be corroborated with achievement in the course. As it stands, the pre-and posttests given in this literature review are not necessarily like assessments students would normally take, and most of the studies reviewed in this thesis lasted a week.

When motivation was investigated in the studies documented in Chapter Two, researchers used self-reports, scored on a Likert scale, as pretests to establish baseline motivation levels. In addition to the limitations of self-reports as outlined earlier in Chapter three, motivation is difficult to measure objectively. Researchers ought to continue to gather more objective methods of quantifying motivation. Nevertheless, new techniques have been developed, such as the method Hodgson et al. (2017) used during a study about engagement in flipped classrooms. First, off-task and on-task behaviors were defined by four or five observable traits each. The behaviors of a subset of students participating in the study were observed and counted by the research team in five-minute intervals. The total number of times the selected students were off-task was subtracted from the number of times they were on task to create a score. Every five minutes a new score was created and eventually the scores were averaged for increased accuracy (Hodgson et al., 2017).

Conclusion

Gamified learning is by no means a guaranteed solution to the motivation and achievement deficits of students. Limitations of self-reports, quasi-experimental studies, sample size, and diagnosing the conditions for SDT affect the ability of educators to generalize the results to all classrooms. On the other hand, there was no evidence in the reviewed studies to suggest that gamification in math classrooms is detrimental to student achievement. Some evidence suggests that students who are already highly motivated may be less motivated after playing games. Most of the evidence points to a positive impact on motivation and achievement in the short term, but the absence of long-term studies leaves the future of gamification in education uncertain.

References

Aloe, A. M., Becker, B. J., Duvendack, M., Valentine, J. C., Shemilt, I., & Waddington, H. (2017). Quasi-experimental study designs series—paper 9: collecting data from quasi-experimental studies. *Journal of Clinical Epidemiology*, *89*, 77–83. https://doi.org/10.1016/j.jclinepi.2017.02.013

Albuquerque, J., Bittencourt, I. I., Coelho, J. A. P. M., & Silva, A. P. (2017). Does gender stereotype threat in gamified educational environments cause anxiety? An experimental study. *Computers and Education*, *115*(1) , 161–170. https://doi.org/10.1016/j.compedu.2017.08.005

American Psychological Association. (n.d.). *American Psychological Association Dictionary of Psychology*. Retrieved April 18, 2023, from https://dictionary.apa.org/self-report-bias

Azuma, T. (1997). A Survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments*, *6*(4), 355–385. https://doi.org/10.1162/pres.1997.6.4.355

Burmeister, E., & Aitken, L. M. (2012). Sample size: How many is enough? *Australian Critical Care*, *25*(4), 271–274. https://doi.org/10.1016/j.aucc.2012.07.002

Caponetto, I., Earp, J., & Ott, M. (2014). Gamification and education: a literature review. In *8th European Conference on Games Based Learning* (pp. 50–57).

Codish, D., & Ravid, G. (2014). Academic course gamification: the art of perceived playfulness. *Interdisciplinary Journal of E-Learning and Learning Objects*, 10, 131–151.

Connell, J. P., & Wellborn, J. G. (1991). Competence, autonomy and relatedness: A motivational analysis of self-system processes. In M. R. Gunnar & L. A. Sroufe (Eds.), *Self processes and development* (pp. 43-77). Lawrence Erlbaum Associates, Inc.

Deci, E. L., Eghrari, H., Patrick, B. C., & Leone, D. R. (1994). Facilitating Internalization: The Self-Determination Theory Perspective. *Journal of Personality*, *62*(1), 119–142. https://doi-org./10.1111/j.1467-6494.1994.tb00797

Deci, E. L., & Ryan, R. M. (2000). The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry*, *11*(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: defining “gamification”. In A. Lugmayr (Ed.), *Proceedings of the 15th International Academic Mindtrek Conference: Envisioning Future Media Environments* (pp. 9–15). https://doi.org/10.1145/2181037.2181040

Dichev, C., & Dicheva, D. (2017). Gamifying education: what is known, what is believed and what remains uncertain: a critical review. *International Journal of Educational Technology in Higher Education*, *14*(9), 1– 36. https://doi.org/10.1186/s41239-017-0042-5

Flunger, B., Hollmann, L., Hornstra, L., & Murayama, K. (2022). It’s more about a lesson than a domain: Lesson-specific autonomy support, motivation, and engagement in math and a second language. *Learning & Instruction*, *77*(1), 101500–. https://doi-org/10.1016/j.learninstruc.2021.101500

Hodgson, T. R., Cunningham, A., McGee, D., Kinne, L. J., & Murphy, T. J. (2017). Assessing Behavioral Engagement in Flipped and Non-Flipped Mathematics Classrooms: Teacher Abilities and Other Potential Factors. *International Journal of Education in Mathematics*, Science and Technology, *5*(4), 248–261.

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. *Grantee Submission*, *67*(1), 423–441. https://doi.org/10.1007/s11423-019-09653-8

Jagušt, T., Botički, I., & So, H.-J. (2018). Examining competitive, collaborative and adaptive gamification in young learners' math learning. *Computers & Education*, *125*(1), 444–457. https://doi-org/10.1016/j.compedu.2018.06.022

Jimenez, K. (2023, April 17). Most Americans are unhappy with the math taught in classrooms, new survey shows. *USA Today*. https://www.msn.com/en-us/news/us/most-americans-are-unhappy-with-the-math-taught-in-classrooms-new-survey-shows/ar-AA19Xkd0#:~:text=A%20mericans%20are%20largely%20unsatisfied%20with%20the%20way,be%20the%20subject%20most%20in%20need%20of%20updating.

Klabbers, J. H. (2018). On the architecture of game science. *Simulation & Gaming*, 49(3), 207–245. https://doi.org/10.1177/1046878118762534

Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: A literature review. *Quality and Quantity*, *47*(4), 2025–2047. https://doi.org/10.1007/s11135-011-9640-9

Landers, R. N. (2014). Developing a Theory of Gamified Learning: Linking Serious Games and Gamification of Learning. *Simulation & Gaming*, *45*(6), 752–768. https://doi.org/10.1177/1046878114563660

Li, S., Shen, Y., Jiao, X., & Cai, S. (2022). Using Augmented Reality to Enhance Students' Representational Fluency: The Case of Linear Functions. *Mathematics (Basel)*, *10*(10), 1718-. https://doi.org/10.3390/math10101718

Mahnič, N. (2014). Gamification of politics: start a new game! *Teorija in Praksa*, *51*(1), 143–161

Manzano-León, A., Camacho-Lazarraga, P., Guerrero, M. A., Guerrero-Puerta, L., Aguilar-Parra, J. M., Trigueros, R., & Alias, A. (2021). Between Level Up and Game Over: A Systematic Literature Review of Gamification in Education. *Sustainability (Basel, Switzerland)*, *13*(4), 2247–. https://doi.org/10.3390/su13042247

Marczewski, A., 2015. User Types. In Even Ninja Monkeys Like to Play: Gamification, Game Thinking & Motivational Design. *CreateSpace Independent Publishing Platform*, 69–84.

McGonigal, J. (2011). Reality is Broken: Why Games Make Us Better and How They Can Change the World. *New York: The Penguin Press*.

Miserandino, M. (1996). Children Who Do Well in School: Individual Differences in Perceived Competence and Autonomy in Above-Average Children. *Journal of Educational Psychology*, *88*(2), 203–214. https://doi.org/10.1037/0022-0663.88.2.203

Niemiec, C. P., & Ryan, R. M. (2009). Autonomy, competence, and relatedness in the classroom: Applying self-determination theory to educational practice. *Theory and Research in Education*, *7*(1), 133–144. https://doi.org/10.1177/1477878509104318

Ottmar, E. R., Landy, D., Goldstone, R., & Weitnauer, E. (2015). Getting from here to there! Testing the effectiveness of an interactive mathematics intervention embedding perceptual learning. *Proceedings of the 37th Annual Conference of the Cognitive Science Society*. 1793-1798. Retrieved February 20, 2020, from https://pc.cogs.indiana.edu/getting-from-here-to-there-testing-the-effectiveness-of-an-interactive-mathematics-intervention-embedding-perceptual-learning/

Park, S., & Kim, S. (2021). Leaderboard Design Principles to Enhance Learning and Motivation in a Gamified Educational Environment: Development Study. *JMIR Serious Games*, *9*(2), e14746–e14746. https://doi.org/10.2196/14746

Rebollo, C., Remolar, I., Rossano, V. *et al.* Multimedia augmented reality game for learning math. *Multimed Tools Appl 81*(1), 14851–14868 (2022). https://doi.org/10.1007/s11042-021-10821-3

Reyssier, S., Hallifax, S., Serna, A., Marty, J.-C., Simonian, S., & Lavoue, E. (2022). The Impact of Game Elements on Learner Motivation: Influence of Initial Motivation and Player Profile. *IEEE Transactions on Learning Technologies*, *15*(1), 42–54. http://doi.org/10.1109/TLT.2022.3153239

Rincon-Flores, E.G., Santos-Guevara, B.N., Martinez-Cardiel, L., Rodriguez-Rodriguez, N.K., Quintana-Cruz, H.A., Matsuura-Sonoda, A. (2023). Gamit! Icing on the Cake for Mathematics Gamification. *Sustainability*, *15*(1), 2334-. https://doi.org/10.3390/su15032334

Rothman, K. J. (2010). Curbing type I and type II errors. *European Journal of Epidemiology*, *25*(4), 223–224. https://doi.org/10.1007/s10654-010-9437-5

Ryan, R. M., & Deci, E. L. (2000). Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *The American Psychologist*, *55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Ryan, R. M., & Grolnick, W. S. (1986). Origins and Pawns in the Classroom: Self-Report and Projective Assessments of Individual Differences in Children's Perceptions. *Journal of Personality and Social Psychology*, *50*(3), 550–558. https://doi.org/10.1037/0022-3514.50.3.550

Ryan, R. M., Stiller, J., & Lynch, J. H. (1994). Representations of Relationships to Teachers, Parents, and Friends as Predictors of Academic Motivation and Self-Esteem. *Journal of Early Adolescence*, *14*(2), 226-249. https://doi.org/10.1177/027243169401400207

Sanmugam, M., Mohamed, H. (2020) Leaderboard: Analysis review behind the usage in gamification. *ASM Sci. J.*, *13*(1), 89–93.

Sánchez-Franco, M. J. (2009): Extrinsic Plus Intrinsic Human Factors Influencing the Web Usage. *In Panayiotis Zaphiris and Chee Siang Ang (ed.), Human Computer Interaction: Concepts, Methodologies, Tools, and Applications*. New York: Information ScienceReference. Retrieved March 22, 2023, from https://www.igi-global.com/chapter/extrinsic-plus-intrinsic-human-factors/22364

Scher, L., Kisker, E., & Dynarski, M. (2015). Designing and Conducting Strong Quasi-Experiments in Education. Version 2. *Decision Information Resources, Inc*. 1-27.

Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action: a survey. *International Journal of Human Computer Studies*, 74, 14–31. https://doi.org/10.1016/j.ijhcs.2014.09.006

Shi, A., Wang, Y., & Ding, N. (2022). The effect of game–based immersive virtual reality learning environment on learning outcomes: designing an intrinsic integrated educational game for pre–class learning. *Interactive Learning Environments*, *30*(4), 721-734. https://doi.org/10.1080/10494820.2019.1681467

Sirakaya, M., & Sirakaya, D. A. (2022). A Case Study: What Attracts Teachers to Augmented Reality. *Participatory Educational Research*, *9*(6), 192–205.

Steele, C. M. (1998). Stereotyping and its threat are real. *American Psychologist*, *53*(6), 680-681. http://dx.doi.org/10.1037/0003-066X.53.6.680

Steinmayr R, Weidinger AF, Schwinger M, Spinath B (2019). The Importance of Students' Motivation for Their Academic Achievement - Replicating and Extending Previous Findings. *Front Psychol*. 2019 Jul 31;10:1730. doi: 10.3389/fpsyg.2019.01730. PMID: 31417459; PMCID: PMC6685139.

Taylor, G., Jungert, T., Mageau, G. A., Schattke, K., Dedic, H., Rosenfield, S., & Koestner, R. (2014). A self-determination theory approach to predicting school achievement over time: the unique role of intrinsic motivation. *Contemporary Educational Psychology*, *39*(4), 342–358. https://doi.org/10.1016/j.cedpsych.2014.08.002

Vallerand, R. J., Pelletier, L. G., Blais, M. R., Brière, N. M., Senecal, C., & Vallieres, E. F. (1992). The Academic Motivation Scale: A measure of intrinsic, extrinsic, and amotivation in education. *Educational and Psychological Measurement, 52*(4), 1003–1017. https://doi.org/10.1177/0013164492052004025

Zhao, J., Hwang, G.-J., Chang, S.-C., Yang, Q., & Nokkaew, A. (2021). Effects of gamified interactive e-books on students' flipped learning performance, motivation, and meta-cognition tendency in a mathematics course. *Educational Technology Research and Development*, *69*(6), 3255–3280. https://doi.org/10.1007/s11423-021-10053-0