2021

Effectiveness of Digital Interactive Experiment in Learning Outcomes and Engagement

Matin Salemirad
University of Central Florida

EFFECTIVENESS OF DIGITAL INTERACTIVE EXPERIMENT IN LEARNING OUTCOMES AND ENGAGEMENT

by

MATIN SALEMIRAD
M.S. University of Central Florida, 2018

A thesis submitted in partial fulfilment of the requirements
for the degree of Master of Fine Arts
in the School of Visual Arts and Design
in the College of Arts and Humanities
at the University of Central Florida
Orlando, Florida

Spring Term
2021

Major Professor: John Thomas Murray

ABSTRACT

Despite some ongoing debates over the positive or negative impact of digital games, educational games are powerful tools to increase engagement and improve learning outcomes. Scientific concepts about the weather are complicated for younger learners and deep learning often requires long-term cultivation. This thesis presents a novel interactive educational simulation called the Science of Meteorology with Interactive Learning Experience (SMILE). In it, students interact with a touchscreen monitor to change weather conditions and learn about clouds and weather science. The relationship between engagement and learning outcomes and the effectiveness of the experience is evaluated using a formal user study. Online data collection was completed after IRB approval during the summer of 2020 with 23 students between the ages of 8 and 12. Student knowledge of clouds and weather science was tested via 12 related questions before playing with SMILE. Furthermore, the impacts of the designed simulation on engagement level were evaluated by the Game Engagement Questionnaire (GEQ), including 19 questions developed by Brockmyer et al. (2009). This study showed a significant improvement in student knowledge where the average of test scores increased 57% for the post-test. The findings show that GEQ needs some modification to measure engagement in different game genres and for different age ranges.

ACKNOWLEDGMENTS

First and foremost, I express my sincere gratitude to E2i creative studio team, and Michael A. Carney for their motivation, enthusiasm, and immense knowledge. I would especially like to thank Eileen Smith, the director of E2i creative studio. As my teacher, adviser, mentor, and ultimate role model, she has taught me more than I could ever give her credit for here. Your insightful feedback pushed me to sharpen my thinking and brought my work to a higher level.

I am truly and deeply indebted to my first supervisor, Professor Joey Fanfarelli. Your friendly guidance, expert advice, and tremendous support have been invaluable throughout all stages of this work.

I cannot express enough thanks to my committee for their continued support and encouragement: Dr. John Murray, my committee chair; Dr. Peter Smith, Dr. Natalie Underberg-Goode and Dr. Maria Harrington. I learned a lot from all of you as your student and also as a colleague.

I would also like to thank Dr. Robert Littlefield, director of the Nicholson School of Communication and Media. You provided me valuable guidance that I needed to choose the right direction and successfully complete my thesis.

Without the support of all members of my family, I would never find the courage to overcome all difficulties during these years. Dr. Mansoureh Kebritchi, Nooshin Vafaie, Asma Agha, and Zahra Bahram: your devotion, unconditional love and support, sense of humor, patience, and optimism are more valuable than you could ever imagine. Finally, I'd like to extend my deepest gratitude to my caring, loving, supportive husband, Rahman. The countless times you kept the children during our hectic schedule will not be forgotten. I wish to thank my wonderful sons, Sadra and Kasra, who provide unending inspiration.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS OR ABBREVIATIONS

Abbreviation	**Explanation**
CL	Collaborative Learning
DGBL	Digital Game-Based Learning
GBL	Game-Based Learning
GEQ	Game Engagement Questionnaire
PhET	Physical Education Technology
SMILE	Science of Meteorology with Interactive Learning Experience
TEL	Technology-Enhanced Learning
WSR	Wilcoxon Signed-Rank

CHAPTER 1: INTRODUCTION

Digital games and simulations are popular in the field of education. Educational games are already integrated into the traditional education process, for better or worse (Carolyn M. Plump, 2017; M.Connolly, A.Boyle, MacArthur, Hainey, & M.Boyle, 2012). Research on the benefits of using educational games and simulations has focused on the impact on learning outcomes, motivation, and engagement. This thesis documents the creation of a novel interactive experience that presents a model of natural phenomena and an evaluation of its impact on learning and engagement (M. O. Riedl & V. Bulitko, 2013). This work builds on and examines related work in serious games, educational games, and educational simulations.

School systems face a lack of student engagement in learning environments. Using educational games and new technologies in the classroom has shown an increase in student engagement (Kim et al., 2017; E. Skinner, Furrer, Marchand, & Kindermann, 2008). Research involving engagement in interactive experiences has also shown a positive impact on outcomes in games with learning systems (Li, Lemieux, Vandermeiden, & Nathoo, 2013). As a result, school systems are considering games as educational applications to enhance engagement and improve learning outcomes (Khan, Ahmad, & Malik, 2017).

Educational Simulations

An educational simulation is a computer simulation created to simplify and accelerate learning (Jimoyiannis, 2009). For over five decades, there has been interest in studying the effects of simulations and games in educational contexts. Educational digital games and simulations have been used for teaching essential skills such as critical thinking, understanding science, and mathematics. However, introducing an exact framework that evaluates the efficiency of games and simulations in particular learning contexts is not easy (Freitas, 2018). In a blended learning method, students learn content using both traditional and online learning. Gamification works best in a blended learning method that is integrating games into non-game contexts. Learning, whether traditional or a blended one with games, is a cognitive process, and if this process does not occur correctly, then the learner will not acquire the knowledge. Klabbers (2009) reported how the usage of many definitions and terminology about different types of games, simulations, and simulation games has led to "terminological ambiguity." Although they all have a broad meaning, in this thesis, "simulation" describes a scenario-based environment to allow students to interact with, learn from, or apply their previous knowledge (Angelini, 2016).

Technology-Enhanced Learning (TEL) is a broad category of techniques and strategies that is used to describe the effects of games and simulations on learning outcomes.TEL describes techniques educators are leveraging such as videos and interactive learning tools in different areas including training, simulations, and health education (Nicoll, MacRury, Woerden, & Smyth, 2018); teacher training (Kenny & Mcdaniel, 2011; Räsänen, Laurillard, Käse, & Aster, 2019); and school education (Hainey, Connolly, Boyle, Wilson, & Razak, 2016). Through TEL, teachers can find more creative and engaging ways to teach science (Eleni A. Kyza & Tiberghien, 2009). Educational games and simulations can be powerful tools to teach science (Sengupta & Clark, 2016). This study aims to demonstrate one potential application of TEL using an interactive simulation to teach weather science.

Engagement and Learning Outcomes from Simulations / Games

Several studies investigated the effects of serious games and simulation on learning engagement. Quinn (2005) reviewed how to make learning more efficient. He used the term “engagement” for any situation where the learner has fully focused attention on a particular task (Quinn, 2005). Student disengagement is also a serious and pervasive problem. Based on the work done by the Organization for Economic Co-operation and Development (OECD), one of the most important challenges facing the education system is student disengagement at schools (OECD, 2016). They define disengagement as a stage that students withdraw from the school because they are not engaged to the learning process and feel that they do not belong to school. Disengaged students do not cognitively involve in learning. Their latest assessment in 2015 focused on science, mathematics, and reading. Regarding science learning, they wrote about the need for students to develop an understanding of science and “think as scientist” through being able to conclude by evaluating evidence. To help students get to this point, education systems should use interactive tools and serious games (Gurria, 2015). Based on research on how students learn when using Physical Education Technology (PhET) interactive simulations, students are more easily engaged and explore science topics with unfamiliar science concepts (Honey & Hilton, 2011; PhET, 2019). Overall, the research has shown that simulations are a promising area of research. SMILE, described in Chapter 3, is an example of an interactive simulation targeting young students to learn about weather science. .

Learning and Engagement

One of the reasons many students prefer playing games rather than studying is because studying can be tedious. Studying that relies on extrinsic motivation are weak and may be detrimental in the long-term (Bénabou & Tirole, 2003). Adding game activities in studying is effective because the desire to interact involves intrinsic, rather than extrinsic, motivation Engaged students experience intrinsic motivation (Adams, Reid, et al., 2008) Game-Based Learning (GBL) provides opportunities for students to have fun achieving learning outcomes.

Research concentrated on using games to enhance motivation, improve engagement, and develop student science understanding. Well-designed digital games proved successful in building a conceptual understanding of science (Lieberman, Fisk, & Biely, 2009) and gain knowledge (Paras, 2005). (Kebritchi, 2008; Linke, Kothe, & Alt, 2017; Mark Prensky, 2003) For understanding science knowledge, engaging in critical thinking, and developing problem- skills, not enough (Beatty & Schweingruber, 2016). Deep learning needs engagement in learning activities. If students are given the chance to apply a simulated phenomenon that they learned theoretically, they will understand how things work in reality, leading them to a deeper understanding of the learning process. The simulation games are structured to present a challenge of a scientific phenomenon and to engage students to solve the challenge (Sengupta & Clark, 2016). In the context of games, there are many ways to measure students' engagement. Game Engagement Questionnaire (GEQ) is a self-report survey to measure engagement at different levels in video game playing (Brockmyer et al., 2009). Find GEQ along with the details about its four constructs, flow, presence, absorption, and immersion in Chapter 2.

SMILE

This thesis presents the design and evaluation of a novel interactive educational simulation of weather conditions for elementary school children. The application, "Science of Meteorology with Interactive Learning Experience" (SMILE), is an interactive digital simulation where students interact with a touchscreen to change weather condition factors and see the results. For example, they can learn that high humidity may affect rain in specific conditions and, if the temperature is low enough, and the conditions where rain changes to snow. For details on the design and development, see Chapter 3.

Design Criteria

Game design is a pedagogical strategy that focuses on broad areas such as introducing a concept, promoting learning, and designing authentic, meaningful, and engaging tools. When discussing game design for learning activities, goals include improving knowledge, motivation, and engagement in the learning process (Vos, Meijden, & Denessen, 2011). Part of this research focuses on identifying which design characteristics help students engage and learn. The design of SMILE was informed by proven research for creating highly engaging and effective learning.

The primary goal of this thesis was to build an educational simulation experience using best practices and design. The secondary goal was to evaluate whether this educational simulation was effective in delivering learning outcomes while maintaining student engagement. Based on the preliminary results, the author believes more interactive experiences like SMILE should be created and evaluated with a larger population.

Statement of the Problem

Millions of people routinely obtain basic information about the weather through apps on their phones, local TV channels, the National Weather Service, or (Lazo, Morss, & Demuth, 2009; Phan, Montz, Curtis, & RiCkenbaCh, 2018)(Lazo, Morss, & Demuth, 2009; Phan, Montz, Curtis, & RiCkenbaCh, 2018). Knowing about the local weather is a crucial topic for many people and this shows in its emphasis in the K-12 science curriculum. In their science courses, students learn about the water cycle, the elements of cloud formation, and other hydrological phenomena. However, past research has shown that early learners often struggle (Barrutia, Ruiz-González, Domingo Villarroel, & Diez, 2019; Maleus, Kilas, & Kruus, 2016; Villarroel &

Ros, 2013). A plethora of studies have also examined the knowledge and misconceptions in different (Zhang, Chen, & Ennis, 2017) science such as physics (Pundak, Liberman, & Shacham, 2017), astronomy (Barke, Hazari, & Yitbarek, 2009), chemistry (Luebeck, Roscoe, Cobbs, Diemert, & Scott, 2017), mathematics (Luebeck, Roscoe, Cobbs, Diemert, & Scott, 2017) for young learners. Moreover, research has shown how students have difficulties learning some conceptual phenomena in physics and earth science (Vosniadou & Skopeliti, 2017) some conceptual phenomena in physics and earth science. Misconceptions in the weather and formation of clouds and precipitation are widely reported (Barrutia et al., 2019; Kikas, Karken, & Malleus, 2017; Salter, Nguyen, & Lankford, 2016). Understanding the weather and cloud formations is one of the earliest abstract concepts to be introduced to children and is related to a wide range of knowledge. Misconceptions become increasingly significant when children build up their knowledge based on missing or incorrect understandings. Thus, it is essential to help children to have a strong conceptual understanding of the science of weather.

Motivation

This thesis presents a educational simulation to help students enjoy the process of learning. There are some games and interactive online activities explaining the formation of clouds and precipitation. However, no published experience focuses on the relationship between humidity and temperature in cloud formation and precipitation in the way presented in SMILE. This thesis documents a novel interactive digital experience relating humidity and temperature at different altitudes to cloud formation. SMILE provides a strong example of the way an educational simulation can keep younger students engaged while learning and having fun. This thesis presents an interactive experience for elementary students and evaluates the changes in their understanding of the associated weather concepts. It also measures engagement as a principal factor in learning.

Hypothesis and Research Questions

This research study focuses primarily on quantitative data, supplemented with questionnaires, and surveys. Pre- and post-test surveys check participant's knowledge before and after interacting with SMILE. Their score represents the participant knowledge gains on the subject of weather and cloud topics after interacting with SMILE.

The following research questions examine the relationship, if any, between the impact of the designed simulation on student engagement as measured by Game Engagement Questionnaire (GEQ) and individual learning achievement. Correlation testing is used to address two research questions and two hypotheses.

Research question 1: Does SMILE produce any statistically significant impact on student engagement as measured by GEQ?

Research question 2: Is there any statistically significant impact by SMILE on student learning achievement?

Hypothesis 1: There is a statistically significant improvement in learning outcomes from pre-test to post-test.

Hypothesis 2: There is a positive correlation between learner engagement, as measured by GEQ, and the knowledge gains of a student in the topic after interacting with SMILE.

Assumptions, Limitations, Scope and Delimitations

This study examined the effects of the specific simulation, designed by the author, on student engagement and learning achievement. We collected survey data from fourth and fifth-grade elementary school students in Florida in the summer of 2020. Based on limitations by the nature of the quasi-experimental research, the results do not stand up to precise statistical scrutiny. This study's primary assumption was based on how students are equal in terms of their engagement before starting the interactive experiment. Consequently, this study's findings do not generalize, but instead provide promising preliminary evidence. A follow-up project could employ the methods described in this thesis in a future study with a broader audience.

Chapter Summary

Chapter 1 introduces the study, the statement of the problem, and the purpose and significance. Chapter 2 contains a review of current research examining simulations and interactive experiences to enhance student engagement, focusing on educational experiences that help students learn more effectively while also enjoying interacting with simulations. The review of existing literature looks at the factors that influence engagement and the ways to measure it. Examples of similar simulations on related topics are also included. Chapter 2 details the artifact, SMILE, and the process used to develop and prototype it. Chapter 3 explains all design aspects of SMILE. Chapter 4 details the evaluation of SMILE, including the questionnaire design, participants, pilot study, and the test plan with its results and limitations. It also mentions the preference of this design over similar games and simulations in the same field. Chapter 5 reports the results and discusses the hypothesis and research questions in this study. The chapter ends with the implications of the thesis findings and suggestions for future research.

CHAPTER 2: LITERATURE REVIEW

Identifying games and simulations used in education are complex. There are many ways to classify games as educational, serious, and entertainment. Serious games are increasingly being used in education. Interest in studying applications for serious games in education has increased in the past decade alongside the increase in the popularity of digital games in society. Research about the effectiveness of game-based learning (GBL) is on various domains such as math (Kim et al., 2017), science (Aghababyan, 2014), and biology (Sadler, Romine, Stuart, & Merle-Johnson, 2013). Although there is abundant research pointing out how game-based learning increases student motivation, some studies have failed to find positive effects of game-based learning and student achievement (Chu, 2014). This chapter focuses on the importance of games in education and game-based learning and reviews research about game design in educational simulations.

Serious Games in Education

Zimmerman and Salen defined games as consisting of both rule-based play and voluntary activity (Salen & Zimmerman, 2004). The application of games to other domains continues to grow in tandem with interest in learning games and educators use them as a tool to enhance student academic performance and engagement (Gee, 2008). Despite early research focusing on the negative impact of computer games such as depression, addiction, and violence (Merhi, Faugloire, Flanagan, & Stoffregen, 2007), several researchers have examined the positive potential of gaming on users (Connolly, A. Boyle, MacArthur, Hainey, & Boyle, 2012) such as supporting learning activity (Foster, Sheridan, Irish, & Frost, 2012), encouraging participation in the class (Denny, 2013), and increasing student motivation to learn (Domínguez et al., 2013).

One type of serious game is designed to deliver educational content in a game-based environment. Abt (1970)established the term serious game in his book of the same title. He defines a serious game as a play activity with an educational purpose, not just amusement. Since then, there have been many changes in game design and development. The current definition of serious games appears to follow the lead set by Michael and Chen (2005) of any game without a primary purpose of entertainment, enjoyment, or fun. Suits (2014) defines playing a game as engaging in an activity with some rules. This perspective is also popular with game developers and educational content developers as well. This thesis adopts Suits definition of game as SMILE is designed to engage players in a rule-based activity, even if it does not include formal win conditions. SMILE could also be considered a casual creator application, as the interactions with the experience are intrinsic and exploratory, though the author focuses primarily on the relationship of the design and evaluation to educational and serious games as they are more common in the literature.

Game-Based Learning

Game-based learning (GBL) is a type of game play aimed at a learning context. Prensky (Marc Prensky, 2003) uses the term “digital game-based learning” to define a game combined with curricular contents. It is designed to teach learners knowledge they can apply in real-world scenarios. The challenge and engagement aspects of games, in particular, can have a significant effect on learning (Hamari et al., 2016). Whether it is fully or partially integrated into a classroom, game-based technologies are a useful tool to support teaching, learning, and evaluation. Students learn faster when they are active participants (Carolyn M. Plump, 2017; C.-C. Chang, Liang, Chou, & Lina, 2017).

Games add a fun factor in learning environments and can significantly improve learning performances. This fun factor tends to motivate and encourage learners to participate in the learning process. Koster (2013), in *A Theory of Fun*, writes, “It is the act of solving puzzles that makes games fun. In other words, with games, learning is the drug” (Koster, 2013). Following this logic, as long as the players' brain is busy with the game, they will desire to continue the game. This logic is also applicable beyond games to other forms of interactive applications. If an interactive experience can keep students engaged, it will encourage the learner to continue learning.

Some studies have reported a direct link between the fun factor and learning (Connolly et al., 2012). In a study on *Kahoot*, a game-based quiz platform, Wang (2015) showed that most students agreed that the fun factor motivates them to learn. Other researchers have focused on adding fun factors in serious games. They believe having more fun promotes motivation, and the outcome will naturally enhance learning performance (Shi & Shih, 2015). There are two fundamental factors of interaction in serious games: fun or pleasure and learning. In other words, serious games can bring fun to the classrooms, and the result is more learning in a fun environment (Huynh-Kim-Bang, Wisdom, & Labat, 2010). Game design patterns are practical design choices that are applicable in all types of games (Bjork & Holopainen, 2004). An effective educational game promotes learning as an educational aspect and is fun to play as a game aspect. Implementing a methodological framework for developing these games is a bridge between the learning and teaching functions and game design patterns (Flores, Paiva, & Cruz, 2020). In some patterns, the learners may have fun at the end of the activity but not learn what was expected. This problem appears when the designer focuses only on interactive elements to motivate users. Some marketing games are good examples in this category such as Bejeweled, Among us, and Overcooked. The solution is using a game-based learning blend pattern where the user learns more than a simple, fun message (Ecker, Müller, & Zylka, 2011). A game-based learning tool must address specific educational objectives. Previous studies have shown that the average student's grade in a game-based environment is higher than the non-game-based one, and students often learn more and faster (Hess & Gunter, 2013).

Designing digital interactive learning experiences means creating a balance between learning and engagement. Too much fun or too little learning and the educational tool is not successful. On the other hand, serious education can be tedious, especially for elementary age students, and it is often difficult to capture their attention for long periods of time. However, since children often spend hours on games, using them to help students concentrate can be useful. Students also do not like playing games where learning is the explicit goal. This is why having a design framework for serious games or simulations is critical in creating successful educational games (Rooney, 2012; Zhu, Pei, & Shang, 2017).

Serious Game Design

How to design serious games is a topic worthy of discussion. Many researchers aim to connect the dots between games and learning. Among many conduction research in this area, the author focuses on the deep learning properties of good games by Gee (2009). Gee defines six properties of good digital game design for promoting learning. The first property focuses on how goals and problem-solving systems give players emotional attachment to the games they play. In general, players want to explore a game's rules to accomplish goals. From this point of view, games are about problem-solving and often self-interest. Some games also focus on winning or losing. In these types of games, players get personally and emotionally committed to accomplishing their goals. This is also a useful feature in serious games where they can encourage cognitive flexibility in a risk-free decision-making environment. Serious games can also use this property in a design focused on real-world or daily events (McGonigal, 2011). The second property defined by Gee (2009) concerns how micro-control creates a feeling of influence. Affecting the movement, action, or any other element in the game is categorized as micro-control for the player. Considerable research on the brain has shown how micro-control can have a stimulating and important effect on humans (Clark, 1997). Gee's third property involves experimental learning opportunities. Through experiencing a situation in a game, a player can be better prepared for future problem-solving outside of it. However, the experiences players have through the game have to be meaningful and allow reflection. A useful experience is one where the player can consider the experience and its context through thinking through the reasons it happened. Good games also help learners to think strategically and give feedback during each step of the experience. The fourth property by Gee (2009) is how player ability must match the actions they are performing. Players should feel challenged by some aspect of the game they are playing. The fifth property is the need for the game to model a combination of the real-world and how parts are simulated. It is crucial for learning to allow a player to view and understand the relationship between the abstract nature of the model and how it is simulated in the game. Finally, the sixth property from Gee (2009) is allowing a player to devise and enact different strategies in the game. This matches the work done by (Shute & Ke, 2012) in examining how embedded assessment can work alongside a player and adjust the difficulty to match the strategies the player develops as they play. The author implements these properties powerfully into the design of SMILE.

Yerkes and Dodson (1908) created an inverted-U model to describe the relationship between pressure and performance as it influences engagement. According to their model

(**Error! Reference source not found.**), the best performance is achieved when people experience a moderate difficulty level. Too much or too little pressure negatively affects performance and engagement. (Schell, 2008) has used this model in his book *The Art of Game Design* to describe the relationship between game difficulty, too. He discusses how players should feel a similar amount of pressure (difficulty) in completing tasks in a game.

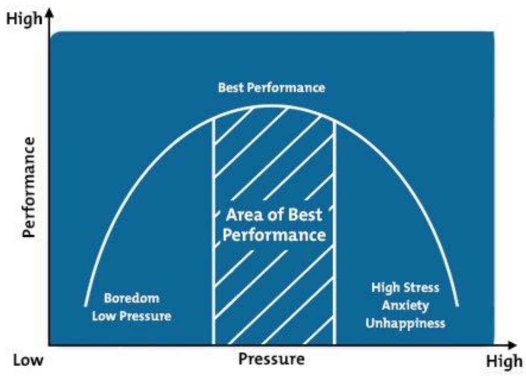


Figure 1: Inverted-U Model. Retrieved from "The Relation of Strength of Stimulus to Rapidity of Habit-Formation" by Robert Yerkes and John Dodson. Published in the Journal of Comparative Neurology (1908). Work now in the public domain

Using Touch Devices as Educational Tools

More young children are using touchscreen devices than ever before. Children are increasingly familiar with tactile devices as they are common not only as communication but entertainment devices in households. Multiple studies have also shown the potential of mobile technology and touchscreen devices for K-12 education (Derick Kiger, 2012; Jenny Eppard, 2016; Min Liu, 2014). Touchscreen devices such as iPads, tablets, and smartphones have become very common among US households (Rideout, 2016). Table 1 shows the results of a recent study measuring the time spent per device on each media activity for young children (Rideout, 2017).

Table 1: Average Daily Time Spent on Media Activities. Retrieved from "Evolution of Media Use by Kids Age 8 and Under 2011-2017," 2017

Watching TV/DVDs/videos	1:29
Tablet	:56
Smartphone	:48
Computer	:31
Video game console	:28
iPad/iPod Touch	:27
Handheld gamer	:07
CD player	:04
E-reader	:02
Total screen media	4:36

Touchscreen devices are increasingly used in K-12 classrooms. Science classes often feature some forms of computers or touchscreen devices for greater student accessibility ("National Academies of Sciences, Engineering, and Medicine," 2019; *Simba Information*, 2016). Students may use these tools for classroom learning, developing models and simulations, watching science videos, or playing educational games designed to explain different phenomena. Such devices frequently play a role to enhance performance in specific education areas to improve academic, daily living, social, and communication skills. Although some research has shown negative effects of using some touchscreen devices such as mobile phones on education (Ting, 2012) because of the small screen and limited input options, there is plenty of research to show valuable ways of using these tools to improve collaborative learning (Viloria, Lezama, & Cabrera, 2020). Researchers at Victoria University of Wellington (S. Gould, Gilbert, J. Pike, & J. Menzies, 2018) use big touch screen monitors to transform a traditional anatomy lab into a corroborative learning environment to assist students' learning. Cavus and Ibrahim (2017) described how an interactive application developed for touchscreen devices helped teach English as a second language through a focus on improving learning skills such as vocabulary, pronunciation, listening, and comprehension. Multi-touch tablets also provide children an opportunity to work individually or in a group and share the results with their peers. Tabletop interfaces were used for digital storytelling purposes to enhance collaboration, creativity, and social interaction (Alofs, Theune, & Swartjes, 2011; Alves, 2010). A recent study evaluated the use of a digital interactive globe system in an Earth science class. These findings demonstrated that engaging the user with hands-on-experience enhanced their understanding of the presented concept (Liou, Bhagat, & Chang, 2018). Many schools also allow students to use hand-held touchscreen devices in the classroom as learning tools ("New Schools 2018-19," 2018). This technology is an opportunity for students to engage and interact with what they are learning (Angeles, 2015). Touchscreen monitors are used not just at schools but also in the museums, too (Block et al., 2015; Roberts et al., 2018). Touchscreen devices encourage both children and adults to engage more with scientific concepts.

Fun Element in Design

The fun element in Games-Based Learning (GBL) comes from virtual characters, environments, narratives or storyline of the game, and their related challenges (Jabbar & Felicia, 2015). Providing an interactive learning game environment that can be explored by the player and integrating multimedia assets, such as sounds, graphics, and animations are critical to motivate learning (Ahmad, Rahim, Osman, & Arshad, 2019). Learners also enjoy making mistakes through experiments in a risk-free environment. A beneficial GBL application draws the learner into a virtual environment that looks and feels familiar and relevant. An educational game can be a continuous activity or simulation in real life. Because the simulation resembles reality, the learner engages with it more.

Some studies have reported no relationship between narrative and learning achievement (Huizenga, Admiraal, Akkerman, & Dam, 2009). Filseckera and Hickey (2014) used a narrative scenario in a GBL environment for the game *Quest Atlantis* (QA). It included aspects of learning and motivation in a gaming environment. The user can make a virtual person and interact with other users or mentors in social interactions. QA also acted as a narrative toolkit. However, studies did not show any additional or deeper engagement with the learning progress because of the game's narrative aspect (Filseckera and Hickey, 2014). Another study analyzed the game design of physics games to find a methodology and design guide for making educational games. Echeverria et al. (2012)compared the effects of narrative elements through a comparison of knowledge acquired by two groups of students playing the game with and without the narrative element. They did not find any significant difference in the result between the two groups (Echeverria et al., 2012). However, the definition of narrative is not the same for all projects. Since this study includes an interactive experience, its narrative definition comes from interacting with the learning experience itself. This form of narrative in a digital learning experience comes from the use of storytelling with structure learning. In other words, storytelling can direct the learner to make decisions that shape the consequences of the story (M. O. Riedl & V. J. A. M. Bulitko, 2013). The simulation presented in this study does not use explicit narrative, but it does use choices as part of its overall storytelling as part of its simulations.

Challenge is another aspect of the fun factor that is associated with difficulty. Challenge is a task or a problem that the difficulty of that depends on the player’s skill, motivation, ability and knowledge (Iversen, 2012)The most common focus of GBL studies is an emphasis on the concept of challenge and its effect on understanding of topics (Jabbar & Felicia, 2015). Several research projects have shown how adding a challenge element to a game increases the player's curiosity (Rosas et al., 2003; Tan, Goh, P.Ang, & S.Huan, 2013). However, it is essential to measure how much challenge is needed to encourage the user to continue playing. Although students may engage more initially if they face too much challenge, some may quit playing (Ke & Abras, 2013). Accordingly, the challenge must be clear and meaningful and varied with level difficulty. A quasi-experimental study by Rosas et al. (2003) evaluated the acquisition of reading and basic math knowledge of elementary students using educational games within aligned with the existing curriculum. Their results showed using challenges in GBL produced higher grades due to a better understanding of the subjects (Rosas et al., 2003; Sadler et al., 2013). Another study on elementary students showed how the challenge element has no meaningful effect on student engagement. However, the presence of a reward encouraged them to concentrate more but didn't extend play time(Ronimusac, Kujalab, Tolvanenc, & Lyytinen, 2014). A skill-balanced challenge and rewarding experience give learners a feeling of pleasure and fun.

Csikszentmihalyi (1997) coined the term "flow" to describe phenomenon of people enjoying a task while focusing on it. This concept draws attention to the close relationship between skill and the complexity of the task. Complexity may contribute to player confusion, but can also be a beneficial emotion to foster engagement in learning environments (K. D'Mello & C. Graesser, 2014). Agarwal and Karahanna (2000) examined several methodologies to measure learner emotions through improvements of the design of interactive digital learning environments (IDLE) (Agarwal & Karahanna, 2000). Most of the time, when a learner acquires complex knowledge or attempts a complicated learning task, confusion can occur. Unexpected feedback on a task or a surprising outcome can contribute to this confusion as well. In these scenarios, a learner may give up on the session. Therefore, a designer needs to put the learner in a situation which causes a certain level of confusion, but not too much where they want to stop. A well-designed interactive system, then, should increase the fun factor in GBL and continually draw the student's attention (Ke & Abras, 2013).

An effective interactive learning environment helps learners choose their actions and experience consequences in a low-risk setting. They can revise their plans and see a new result. This kind of activity promotes engagement. There is a correlation between enjoyment and engagement, and studies have proven engagement is more important than the role of fun in serious games (Iten & Petko, 2016). When children are eager to play a serious game, their primary interest is not fun but achievement. The use of interactive games can be an inspiring alternative to increase engagement at schools, too. Games provide a creative environment for teaching knowledge and improving school skills by adding diversity in educational games. They can act as a supplemental element in the classroom or as a partial substitute for traditional education methods (Lameras et al., 2017b; Rutten, Van Joolingen, & Van Der Veen, 2012; Sitzmann, 2011). Research has also indicated how games promote knowledge acquisition and teach through a playful learning environment. In particular, games influence student motivation and engagement in the classroom (Emmanuel Tsekleves, 2016; Tsekleves, Cosmas, & Aggoun, 2016; Whitton, 2007).

Engagement

Engagement has been typically described as having two aspects: behavioral and emotional, sometimes labeled affective(Marks, 2000; Willms, 2003) Engagement has been defined by multiple aspects. In education, engagement can refer to the attention, motivation, interest, and passion. . These have been linked to behavioral, cognitive, and emotional qualities (Fredricks, Blumenfeld, & Paris, 2004; Hoffman & Nadelson, 2010; M.-T. Wang, Willett, &

Eccles, 2011).(Zhu et al., 2017) Many researchers adopted this definition of engagement; however, a limited number of studies have measured multiple domains of engagement concurrently (Filseckera & Hickey, 2014; Gale M. Sinatra, Heddy, & Lombardi, 2015; M. T. Wang & Eccles, 2012). In addition to the three-aspect models, some researchers have gone further than these definitions and added another engagement dimension. Reeve and Tseng (2011) call agentic a fourth type of engagement. Agentic engagement, they describe, occurs when a student actively contributes to the flow of a task. When students request extra information or show more interest in content, agentic engagement is occurring (Reeve & Tseng, 2011). A systematic review of engagement shows that this concept remains a confusing one with several different definitions, many of which can depend on different factors (Bouvier, Lavoué, & Sehaba, 2014; Boyle, Connolly, Hainey, & Boyle, 2012).

Game, Engagement, and Learning

Many students characterize traditional classrooms without games and interactive technologies as boring. Students are more willing to connect with and play games that are active, fun, and present multitasking interfaces (Marc Prensky, 2003). They want engagement, one of the most widely used terms in education and psychology. In the context of educational psychology, “engagement” relates to academic motivation and achievement (Christenson, Reschly, & Wylie, 2012). Sinatra et al. (2015) believe that student engagement is the “holy grail of learning.” Engagement is a necessary term, they explain, because it has a connection with positive learning outcomes in and out of the academic domain (Gale M Sinatra, Heddy, & Doug, 2015). The number of studies on educational game engagement and its effect on learning and motivation are increasing (Hawlitschek & Joeckel, 2017; Jabbar & Felicia, 2015). Some studies defined engagement as a learner’s involvement and interest toward an educational topic that positively correlates with learning outcomes and behavioral change (Perski, Blandford, West, & Michie, 2017). (Bourgonjon, Valcke, Soetaert, & Schellens, 2010) Hawlitschek and Joechel (2017) examined the effects of digital educational games on learning outcomes. Their results show that educational games increase learner mental effort and ability to match the learning outcomes of the experience (Hawlitschek & Joeckel, 2017).

Definition of Engagement

There are many different definitions of engagement in education, psychology, and game studies. Many researchers believe that engagement is the same as motivation (Aghababyan, 2014; Izar, Stark, Trentacosta, & Schultz, 2008). However, some authors distinguish engagement from motivation (Delaney & Royal, 2017; Delialioğlu, 2012). One of the first definitions of academic engagement linked psychological investment with learning (Newmann, Wehlage, & Lamborn, 1992). According to Skinner and Belmond (1993), children who are engaged in a task show positive emotional attitudes and higher grades. They believe engagement includes both behavioral and emotional components (E. A. Skinner & Belmont, 1993). Other studies have defined engagement in a more detailed manner, including the ways excitement and an enjoyable state of mind can make an activity more meaningful and foster a feeling of competence (Rozendaal, Braat, & Wensveen, 2010). Some definitions connect engagement and academic achievement. They mention how the challenge level of a task can determine the level of engagement (Hoffman & Nadelson, 2010).

After analyzing different literature on engagement in educational games and considering the development of the game engagement questionnaire, the author adopted the definition of engagement based on Brockmyer et al. (2009). In this definition, game engagement is equivalent to game involvement, including absorption, flow, presence, and immersion

(Brockmyer et al., 2009). Each of these represents a factor in the game engagement questionnaire (GEQ) used to measure engagement in participants in this study.

Game Engagement Questionnaire

There are multiple ways to measure engagement in video games. Mouse movement and the number of clicks in a game can indicate engagement (Heather L. O'Brien & Toms, 2008). Ravaja et al. (2006) found a meaningful correlation between game events, emotional attention, and related psychophysiological responses by measuring electro-dermal activity (EDA, previously known as Galvanic Skin Response) and facial electromyography (EMG) (Ravaja, Saari, Salminen, Laarni, & Kallinen, 2006). However, EDA and EMG are also used to identify various emotional states such as boredom, frustration, and enjoyment, not just engagement. There are many questionnaires researchers may use depending on their focus area. Three prevailing approaches are Immersive Experience Questionnaire (IEQ), Player Experience of Need Satisfaction (PENS), and Game Engagement Questionnaire (GEQ) (Yang, Rifqi, Marsala, & Pinna, 2018). A fourth, the Engagement Questionnaire (EQ), has 46 items and uses some questionnaires from psychological literature such as the Presence Questionnaire (Witmer & Singer, 1998) to measure five factors in engagement labeled interest, authenticity, curiosity, involvement, and fidelity (Mayes & Cotton, 2001). Game Experience Questionnaire (GEQ) has a modular structure that consists of a core module of 33 questions with others covering in-game, social presence, and post-game modules. It includes seven areas to measure as game experience: immersion, flow, competence, positive and negative affect, tension, and challenge (IJsselsteijn, 2013). Social Presence in Gaming Questionnaire (SPGQ) measures three psychological elements: empathy, negative feelings, and behavioral engagement (De Kort, IJsselsteijn, & Poels, 2007). The Flow State Scale is designed to measure the flow experience based on Csikszentmihalyi's Flow Theory (Csíkszentmihályi, 1997; Jackson, Martin, & Eklund, 2008), while the Dispositional Flow Scale (DFS) assesses the tendency of users to experience flow (Jackson et al., 2008). A newer version of DFS is DFS-2. It has been widely applied in education and digital games (Hamari & Koivisto, 2014). The immersion experience questionnaire, or Game Immersion Questionnaire (GIQ), measures game enjoyment and immersion as an outcome of a positive gaming experience. This questionnaire has sixteen pairs of questions with different wording to control wording effects (Jennett et al., 2008). Some researchers believe flow is an optimal and extreme state. They explain how immersion is a better word to describe experiencing digital game-play (Cheng, She, & Annetta, 2015; Jennett et al., 2008). Others believe immersion is considered as a precondition of flow and not an extreme state. Jennett et al. (2008) write how engagement is the first step of immersion with the second stage as engrossment. This definition overlaps with Brown and Cairns' (2004) study defining immersion as engagement, engrossment, and total immersion. In a digital game-based science learning, GIQ is a good and reliable questionnaire for proving immersion leads players to higher gaming performance (Cheng et al., 2015). The Temple Presence Inventory (TPI) measures presence as engagement, social richness, social, and perceptual realism (Lombard, Weinstein, & Ditton, 2011). Wiebe et al. (2014) devised the User Engagement Scale (UES) for measuring engagement in digital games. They explored a revised scale using four factors rather than six in the original scale. In their research, focused attention, perceived usability, aesthetics, and satisfaction are the main factors that demonstrated better psychometric analysis than the original six subscales (Wiebe, Lamb, Hardy, & Sharek, 2014).

Although there is no single way to measure engagement in a game-play environment, a self-reporting questionnaire is considered a reliable measurement approach. Brockmyer et al. (2009) explain the Game Engagement Questionnaire (GEQ) as providing a "psychometrically strong measure of the type of engagement elicited while playing video games" (Brockmyer et al., 2009). The GEQ is one of the most well-known questionnaires to measure engagement in video games and can be applied to any digital interactive experience (Heather L O'Brien & Toms, 2010; Psaltis, Apostolakis, Dimitropoulos, & Daras, 2018). Brockmyer et al. (2009) created GEQ consists of 19 questions to measure four parts of engagement: absorption, flow, presence, and immersion (Brockmyer et al., 2009). The pilot version of this questionnaire, based on Brockmyer et al. (Brockmyer et al., 2009) research, was administered to children in fourth through sixth grade. Table 2 is a summary of research on young children. The next section explains four parts of engagement measured by GEQ in details.

Table 2: Summary of Research Measuring Game Engagement on Young Children

Author(s), Year	**Sample**	**Method**	**Result summary**
Brockmyer et al., 2009	17, 4th through 6th grade	GEQ	Develop and test questionnaire to measure video game engagement.
Hong et al., 2013	209, 5th and 6th grade	Flow and engagement	Different game methods affect the player's flow experience.
M. Chang, Evans, Kim, Deater-Deckard, & Norton, 2014	100, 5th grade	GEQ	Girls have higher engagement than boys on the educational video games.
Dele-Ajayi, Sanderson, Strachan, & Pickard, 2016	37, 8 to 16 years old	Immersion survey	Combination of questionnaire and interview shows more engaging experience support effective learning.
Vangeel et al., 2017	1016, 14 to 16 years old	GEQ	Engagement is an important determinant of playing any games as well as serious games.
Psaltis et al., 2018	72, 8 to 10 years old	GEQ combine with motion and facial expression	Combination of using real time engagement to measure affective engagement and categorize the data based on GEQ is an effective way to measure game engagement.
Zulehmay et al., 2019	26, 11 Years old	GEQ	Design and development of a solar system game can provide engagement to the students while getting the basic concept of the topic.

Absorption

Most definitions of absorption describe a state of in-depth involvement and attention in an experience (Agarwal & Karahanna, 2000). Absorption can be either negative or positive. Playing violent video games increases absorption, for example, but can have negative impacts on behavior. While absorbed during gameplay, some negative effects players report are feeling being trapped or weird. Positive absorption, on the other hand, is a reaction to gameplay such as fun, liking, and pleasurable (Blythe & Hassenzahl, 2005), although it may include challenges (Belgium, 2010).

Flow Theory

Csíkszentmihályi (1990) described flow as a deep cognition and emotional mental state people experience while in extreme focus on a task. Research shows how GB increases flow experience to enhance motivation and effective learning (C.-C. Chang, Liang, Chou, & Lin, 2017; Cheng et al., 2015). Flow theory is based on challenge and skill and the balance between these two. If the challenge is based on an appropriate skill level, it would be one important factor to increase the flow experience. Having a challenge in GBL is important. However, it must be a factor of the user's skill level. If the challenge makes the experience too difficult, students can be frustrated or disappointed. Conversely, the users will get bored with an effortless GBL experience. Therefore, the experience must be matched with the learner's skill to have a flow experience. When the ratio between perceived skills and level of challenge is large, people may experience considerable anxiety (**Error! Reference source not found.**).

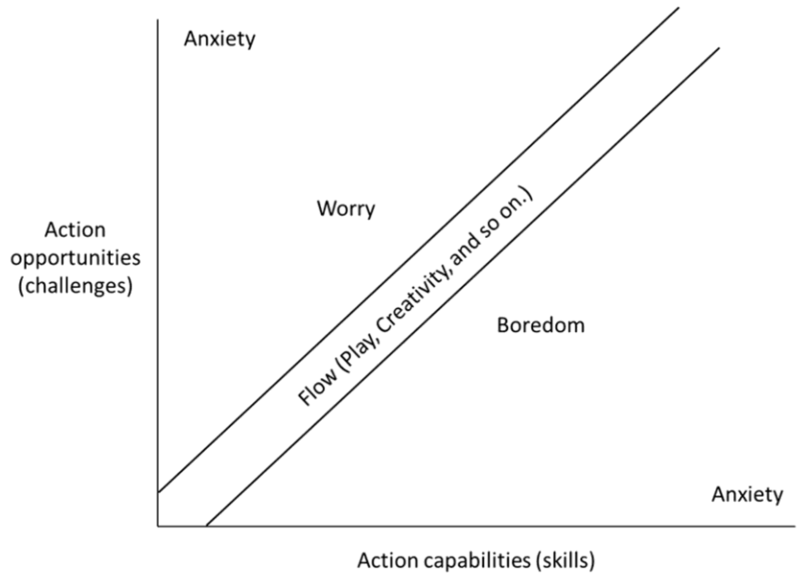


Figure 2: Csikszentmihalyi Flow Model, Adapted from Csikszentmihalyi,1975

Csíkszentmihályi (1990) includes nine dimensions from measuring a flow experience. However, not all researchers apply these factors in their studies. Some only measure flow conditions like challenge-skill balance, clear goals, and immediate feedback (Kiilia, Freitas, Arnab, & Lainema, 2012). In these cases, a sense of control, action awareness, and loss of self-consciousness, time distortion, concentration, and autotelic experience may also occur. Several methods have been used to study flow experience, in which self-reporting techniques are the most common method (Weber, Tamborini, Westcott-Baker, & Kantor, 2009).

Presence

Presence has been understood as a psychological state described as some sort of self-awareness in an immersive experience. According to Draper et al. (1998), presence is a special type of flow experience named telepresence or special presence. In terms of computer games, some researchers believe the two main elements of presence are immersion and involvement

(Weibel & Wissmath, 2011; Witmer & Singer, 1998). The presence questionnaire presented by Witmer and Singer (1998) was developed to measure a virtual environment (VR). As a result, not all questions can be applied to a video game, especially a learning one.

Immersion

There are competing definitions for immersion. Coomans and Timmermans (1997) describe one of common definitions as, "the feeling of being deeply engaged where participants enter a make-believe world as if it is real" (Coomans & Timmermans, 1997). Immersion is an essential experience of interaction. Jennett et al. (2008) measured immersion in videogames based on the three dimensions of game flow, absorption, and presence (Jennett et al., 2008). Brockmyer et al. (2009) added immersion to develop another scale for measuring game engagement.

Questionnaire and Response Measure

It is essential to create an appropriate questionnaire for each design. For young children, creating an enjoyable, easy to read, and easy to answer survey is critical (Guin, Baker, Mechling, & Ruyle, 2012). Simplicity is important for questionnaires involving children. Using short questions with unambiguous and straightforward language is essential (Bell, 2007). The other key is to use positive statements. In essence, a negative questionnaire is not recommended for most surveys. Some researchers believe equal balance in positive and negative statements enhances data quality (Borgers & Hox, 2000).

Each researcher can make different choices as to how they employ the scaling options, such as numerically, visually, and ideographically or a combination these. According to Bell (2007), the best questions should have scales which are clear and easy to perceive for children.

Verbal labels are easier to understand than numeric ones, while visual images have also produced good results. In fact, the respondent will pay more attention if numeric and visual scales are combined with words or pictures to anchor the scale (O'Muircheartaigh, Gaskell, & Wright, 1995). Moreover, responders to the questions with images spend less time than responders to questions with only words or numbers (Stange, Barry, Smyth, & Olson, 2016).

The Visual Analogue Scale (VAS) is frequently used as a psychometric scale to measure the degree of pain experienced by a patient (Hayes & Patterson, 1921). VAS is used to improve the Likert scale in self-reporting measurement designs (Gorrall, Curtis, Little, & Panko, 2016). VAS is an image-based Likert scale anchor without numbers or words. The most common images are smiley faces that range from very sad to very happy faces. Smiley Face Likert (SFL) often appears in UX design for children (Hall, Hume, & Tazzyman, 2016). Further, as many studies report, using colorful and cartoon style emoji to enhance the features is more effective than using just simple black and white smiley faces (Hall et al., 2016; Reynolds-Keefer & Johnson, 2011).

Response Layout

Designing question layout and its effect on rating scales is important. Poorly designed layouts can introduce social desirability or acquiescence bias (Johnson, 2018). One effect the author considered in the current design is left-side bias. This is where survey-takers tends to choose the left-side of the response scale in a horizontal set-up questionnaire. The following example explains this difference in the response order based on the statement “I enjoy learning science in this game.”

Table 3: Ascending-Ordered Response Scale

Strongly Disagree	Disagree	Undecided	Agree	Strongly-Agree

Table 4: Descending-Ordered Response Scale

Strongly-Agree	Agree	Undecided	Disagree	Strongly Disagree

Most students have a slight bias to the item presented first. This is why considering a left-side scale matters (Nicholls, Orr, Okubo, & Loftus, 2006). Although no one ordering or another is completely right or wrong, based on Sauro's (2010) work, "a dishonest researcher who wants responses to be slightly higher in an agreement can place the favorable response options on the left" (Sauro, 2010). This can also create slightly noticeable difference when the sample size is larger than 100 responses, which is not the case in this study. To avoid the left-side bias, we decided to display the options from negative to positive in a horizontal setup.

Level of Guidance, Questions and Clues

A game designer's goal is to have players become self-directed, self-motivated learners (Garris, Ahlers, & Driskell, 2002). Moreno and Mayer (2005)state how learning through video games without any guidance would be inefficient. Their results of three different experiences support appropriate use of guidance as necessary for interactive digital games (Moreno & Mayer, 2005). In research studies on Physical Education Technology (PhET) projects, four levels of guidance are defined (PhET, 2019). These detail the level of simulation: no, driving, gently guiding, or strongly guiding (Adams, Paulson, & Wieman, 2008). In the first and second types, the quality of simulation can strongly affect user engagement and learning. Games without any instructions can intimidate students and prevent them from engaging. The driving questions provide direction for students to explore deeply and build a mental framework. This type is useful for those types of simulations needing more attention to a specific concept. However, comparing with the first type, learners may explore fewer features, although they may explore other concepts. Gently guided simulations include a series of questions in a specific bar or windows on the screen. Strongly guided is a type of instruction that is considered like a “cookbook.” It explains all concepts, sliders, and options in the simulation. It shows all steps learners need to know while interacting with a learning game or simulation. In this case, acquiring knowledge is minimal. Too much information is not helpful as well. It limits the student’s interaction with the simulation with them asking questions and exploring without direct answers. The result of extensive simulation interviews and studies show driving questions and gently guided are the most effective ways to guide the player. These types of guidance create a mental framework for players to approach problem solving and knowledge acquisition (Adams et al., 2008; Aldrich, 2009; Schwartz, Bransford, & Sears, 2005). This must also match complexity level. It must be appropriate for the participant’s age and general knowledge. Likewise, the level of guidance in the simulation affects their understanding and level of engagement (Akaygun & Jones, 2014).

The right design should foster engagement and motivation to learn. The current study uses guidance to instruct players to explore via driving questions. Pesare et al. (2016) state that learning games are successful if they help students build on previously learned knowledge. In a version of the *Handbook of Motivation at School,* the authors mention that If the topic includes relatively unfamiliar science, students can engage effortlessly (Wentzel & Miele, 2009). SMILE uses information fourth and 5th graders already know, such as the meaning of temperature, humidity, altitude, and the concept of forming clouds. One effective way to encourage exploration and engagement is to include some questions and add some clues related to the experience's learning goals (Lameras et al., 2017b). Some of the questions designed for SMILE are easily answered by interacting briefly with the experience. Other questions require investigating the simulation.

Importance of Clouds and Learning Science of Meteorology

Many organizations prioritize teaching meteorology. The World Climate Research Program (WCRP) sponsored by the World Meteorological Organization (WMO), National Aeronautics and Space Administration (NASA), European Space Agency (ESA), and many other centers have several missions, conferences, and workshops related to better understanding of clouds, precipitation, climate, and weather forecasting as well as developing and improving models and simulation on components in the Earth system ("Earth CARE - Earth Online," ; "NASA - ACE," ; "UCP 2019: Understanding Clouds and Precipitation | GEWEX Events," 2019). In Florida, the Next Generation Sunshine State Standards (NGSSS) describes the knowledge or ability students should be able to demonstrate at the end of each grade. Earth and space science appear in 29% of the content in fourth and fifth-grade science textbooks. Students in fourth grade learn about the weather and its types. In fifth grade, students learn about weather patterns, water cycle, and precipitation (NGSSS, 2019).

Examples of Games about Clouds

Despite a growing interest in designing serious and educational games, there are few science simulations about clouds and weather. As educational games become more widely explored, it is important to pay special attention to the design, included concepts, and target groups of serious games and simulations. There are many resources explaining the cloud categories and characteristics of each. However, digital games have an unexplored potential to explain this interesting meteorology fact to younger learners. Some apps or simulations explain how clouds are made and how they change at different altitudes. This section provides examples of games about weather, clouds, and meteorology for younger learners.

Weather Maker is an online interactive weather maker simulation. It explains weather factors such as air pressure, temperature, winds, and humidity. It is published on the Scholastic.com website. As shown in Figure 3: Weather Maker Screenshot. Retrieved from http://teacher.scholastic.com. Screenshot by author, this simulator provides a good understanding of changing temperatures towards the poles and equator. It also includes an explanation of the weather under the image. However, it does not explain clouds, their shapes, or the effect of elevation effect on the weather.

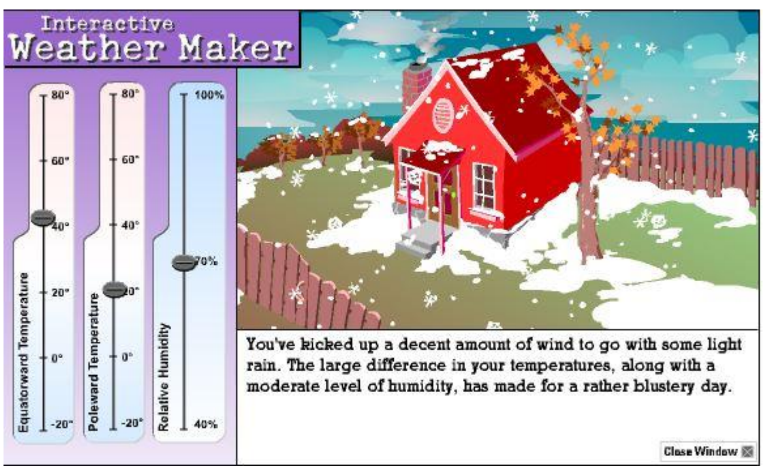


Figure 3: Weather Maker Screenshot. Retrieved from http://teacher.scholastic.com. Screenshot by author

Explorelearning.com is a website with interactive science simulations for grades 3 to 12. It provides an online learning experience aligned with the latest educational standards. One of their gizmos is "Comparing Climates." Students can visit different virtual locations around the world and compare their local climate and landscape through temperature, humidity, rainfall, and wind speed data (Figure 4: Comparing Climates Screenshot. Retrieved from https://www.explorelearning.com/index.cfm?method=cResource.dspView&ResourceID=1080). Screenshot by author).

Figure 4: Comparing Climates Screenshot. Retrieved from https://www.explorelearning.com/index.cfm?method=cResource.dspView&ResourceID=1080). Screenshot by author

The University Corporation for Atmospheric Research (UCAR) Center for Science and Education website has multiple games and simulations related to meteorology with focuses on different topics and learning goals. Their "Make a Thunderstorm" simulation presents different temperatures and humanity options to the player to create a thunderstorm. Figure 5: Make a Thunderstorm Screenshot. Retrieved from https://scied.ucar.edu/interactive/make-thunderstorm. Screenshot by author shows how the game presents a short panel to explain instructions and asks the player questions about the experiments. In their "Cloud Sorting Game" (Figure 6: Cloud Sorting Game Screenshot. Retrieved from https://scied.ucar.edu/interactive/cloud-sorting-game.

Screenshot by author), the player is presented with photos and asked to match the names of different types of clouds by altitude. Unlike many other games, the cloud sorting uses real pictures of clouds.

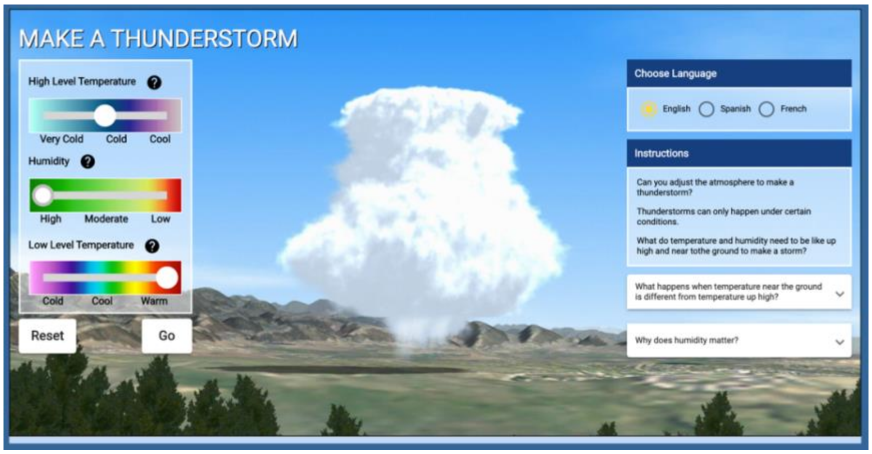


Figure 5: Make a Thunderstorm Screenshot. Retrieved from https://scied.ucar.edu/interactive/make-thunderstorm. Screenshot by author

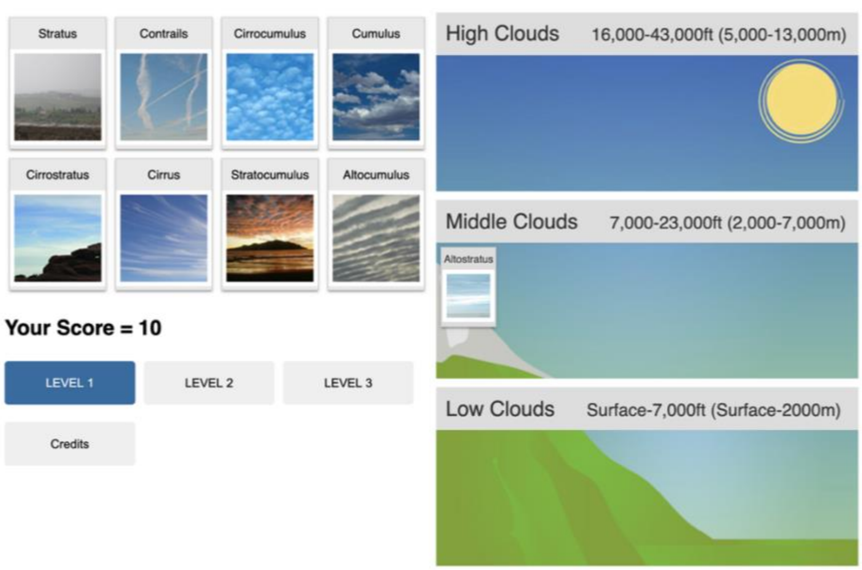


Figure 6: Cloud Sorting Game Screenshot. Retrieved from https://scied.ucar.edu/interactive/cloud-sorting-game. Screenshot by author

The Cooperative Institute for Meteorological Satellite Studies (CIMSS) provides multiple educational projects on meteorology. “Type of Precipitation” (Figure 7: Type of Precipitations Screenshot. Retrieved from http://profhorn.meteor.wisc.edu/wxwise/. Screenshot by author) is an applet by Whittaker and Ackerman (2019) that presents an interactive simulation to explore types of precipitation based on the temperature of the air and dew point. Data is presented in a Skew-T, a useful meteorological thermodynamic chart that presents weather factors such as water vapor, temperature, and humidity below and above 16 kilometers above sea level ("Skew-T Log-P Diagrams," 2019; "UCAR,").

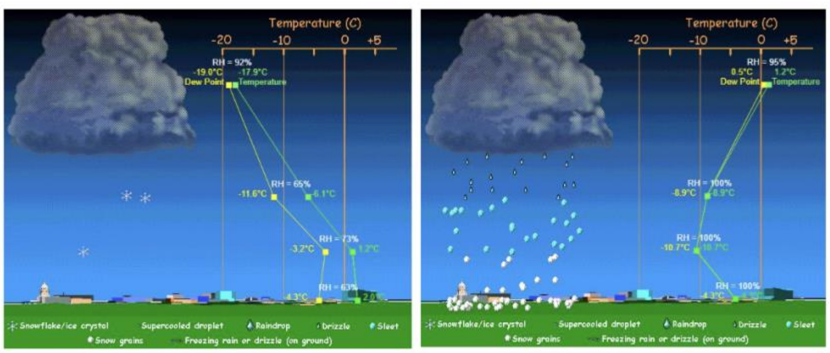


Figure 7: Type of Precipitations Screenshot. Retrieved from http://profhorn.meteor.wisc.edu/wxwise/. Screenshot by author

“Thickness of Clouds” is another project by CIMSS. This simulation helps learners to modify surface temperature and cloud thickness from a satellite view. To start engaging with the learners, this app provides users questions like “What happens to the image when you make the cloud thicker?” and “What happened to the image when you make the cloud thinner?

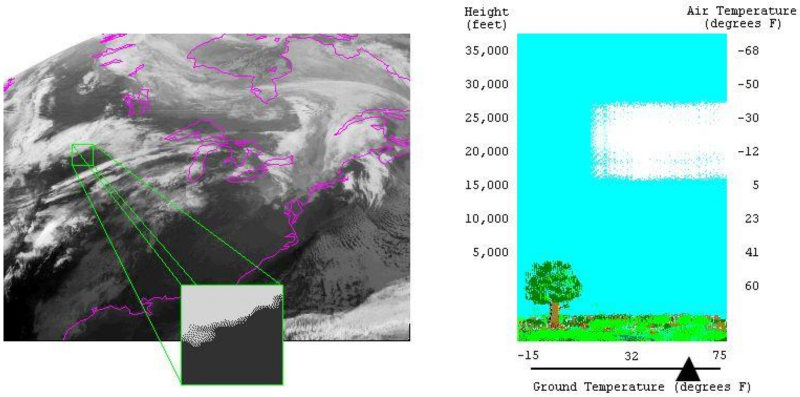


Figure 8: "Thickness of Clouds". Retrieved from http://profhorn.meteor.wisc.edu/wxwise/satir/h5/irthick.html. Screenshot by author

The final simulation presented by CIMSS covers making a shower, fog, or thunderstorm based on the dew point, pressure, and temperature (Figure 9: "Make a shower or a thunderstorm!". Retrieved from http://profhorn.meteor.wisc.edu/wxwise/thermo/h5/tstm.html. Screenshot by). The player can change the stability of the atmosphere through dragging atmosphere sounding or profiling. It illustrates simple images of clouds and rain.

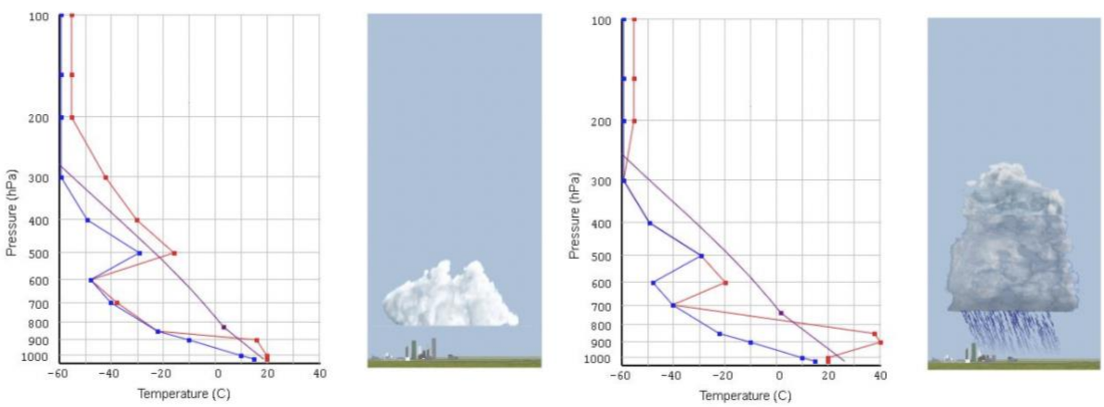


Figure 9: "Make a shower or a thunderstorm!". Retrieved from http://profhorn.meteor.wisc.edu/wxwise/thermo/h5/tstm.html. Screenshot by author

The following chapter details the design of SMILE and game design elements in depth. It also includes different screenshots of the simulation.

CHAPTER 3: THE DESIGN OF SMILE

This chapter describes the educational elements and game design art and theories of creating SMILE. Design elements include goals, interaction, freedom, challenge, and interactive problem-solving. *The Art of Game Design* (Schell, 2008) defines game design from different perspectives or lenses. Based on game design elements by Mifrah Ahmed (2019) the lenses described by Jess Schell (2008), this chapter documents the design process followed in the creation of the science of meteorology in dept in Appendix A.

The process of game design is a complex task. This chapter intends to explain the design of SMILE in details based on the recent and prominent game design elements. Figure 10 is a selected part of a proposed categorization of game design elements adopted from M. Ahmad (2019). The key design elements are grouped into five fundamental categories, namely, mechanics, aesthetics, pedagogy, story, technology

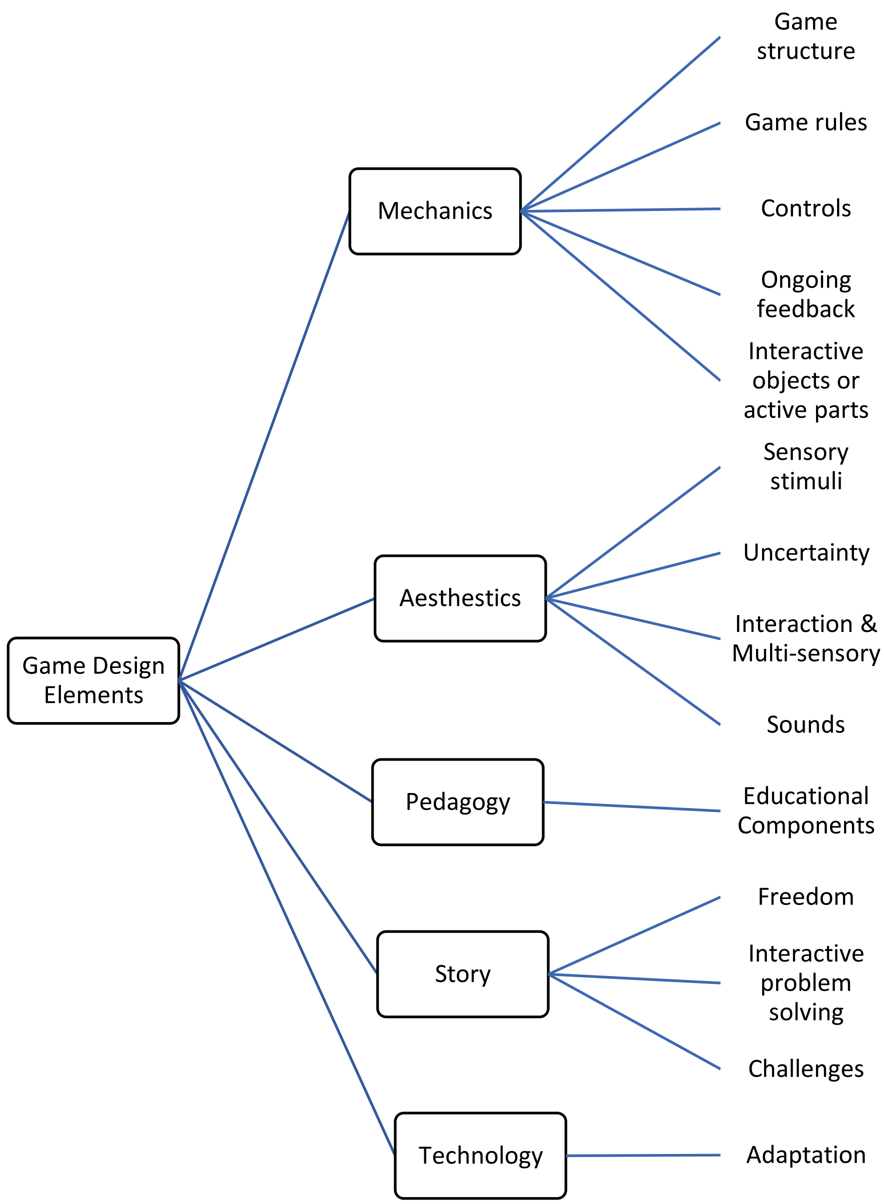


Figure 10: Part of a proposed categorization of game design elements. Adopted form "Game Design, and Intelligent Interaction."

Interaction with SMILE starts with a short tutorial (Figure 11). The tutorial orients the user to the various controls and objectives.

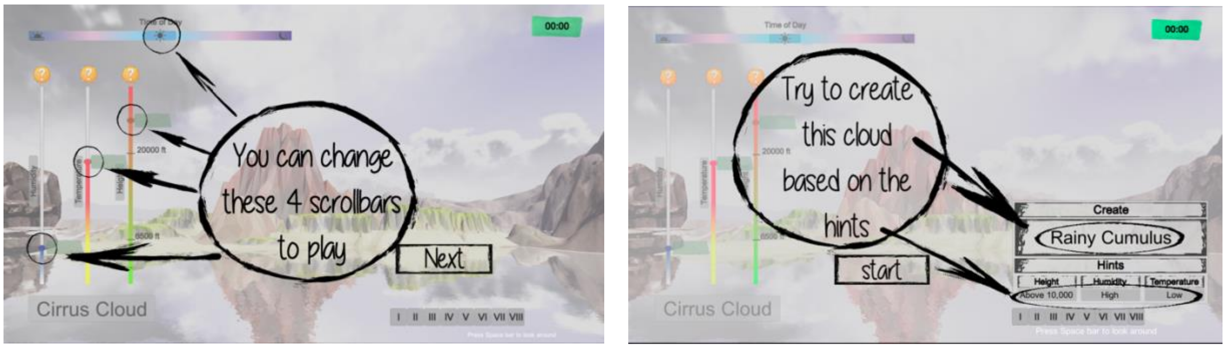


Figure 11: Two screenshots of SMILE tutorial after the welcome screen (Figure 16)

After the tutorial section, a panel on the bottom left side of the windows asks the player to create different cloud every 30 seconds by following the hints. It is not necessary for the players to follow the hints. They can change the sliders for height, temperature, humidity and time of day and explore different scenes to create diverse clouds and make precipitation. Find the information of clouds and their combinations in Appendix A. The final minute of the interactive experience asks students to explore the simulation without explicit guidelines (Figure 12)

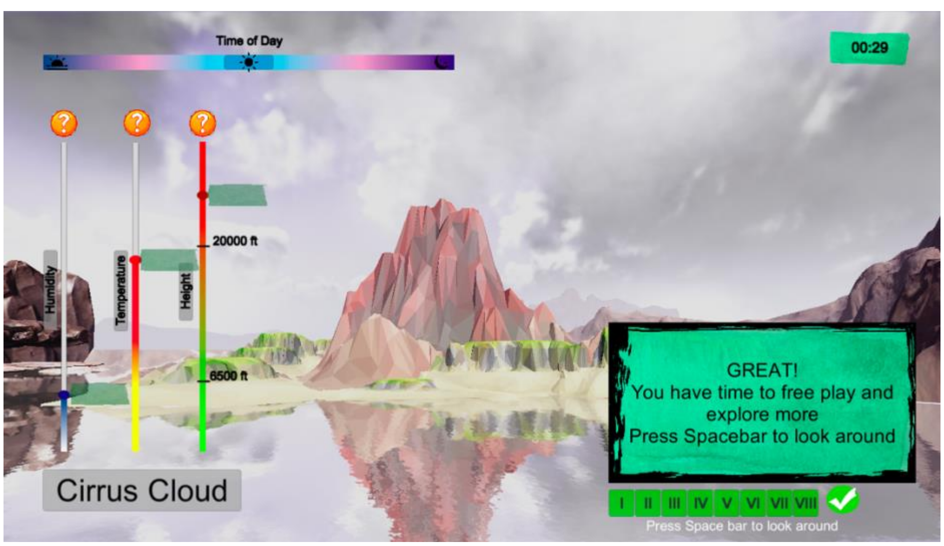


Figure 12: Screenshot of SMILE, final 30 seconds of interaction

Mechanics

The SMILE’s structure includes one main scene after two panels of introductions. The goal is to teach students about clouds and weather through an interactive educational simulation. The supervised part of the game guides players in creating different type of clouds. A new hint is provided every 30 seconds. The hint panel presents the necessary elements to produce each specific type of cloud. Even though these hint panels help players, they are free to explore the game environment and ignore the hint panels’ direction.

Based on Figure 10, feedback and interactive objects are under mechanics. Ongoing feedback improves the learning experiences' outcome. The lens of control (Schell, 2008) describes how important an intuitive interface is to give the player a feeling of control. Each slider in the interface shows an immediate change in the simulation scene. There are three sliders that control humidity, elevation and temperature. It gives players a feeling of influence and controls the game.

The interest curve lens led me to start the experience with a simple challenge (Schell, 2008). For example, the first cloud is cirrus, which is a very common cloud in the sky. The players do not need to change any sliders and the first 30 seconds is for them to get more familiar with the game environment. The second challenge is just to change the humidity slider to see its effect on the density of clouds. Figure 13 a and b show these two challenges. The create panel raises interest gradually by asking for rainy and snowy clouds. As explained in the result chapter, players were interested in the second minute when they could make snow and in the third minute when they saw the rainbow after a rainy scene. Up to the end of the interactive experience, players learned about clouds types based on different heights and possibility of rainfall or snow by changing the temperature and humidity. The game's programming logic follows the formula that creates precipitation based on the relationship of temperature and humidity. Please see Appendix B for more information about the science behind the game logic.

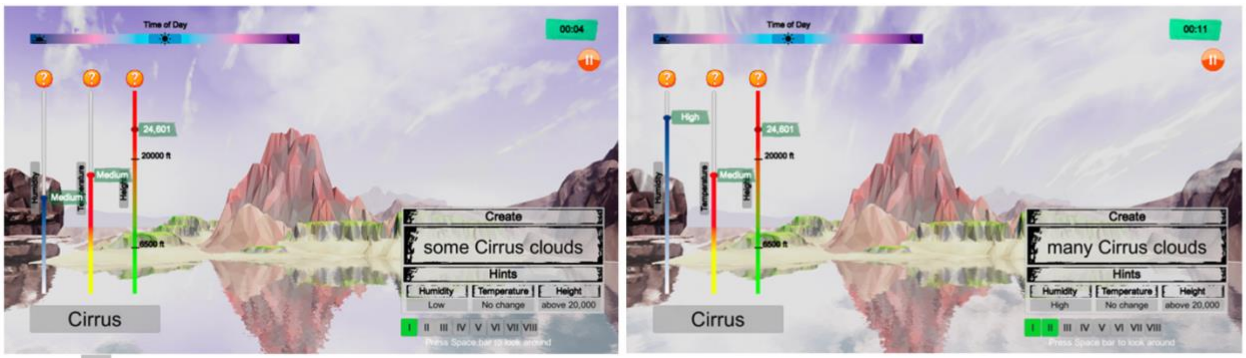


Figure 13: Sample scenes of cirrus clouds

While players can indirectly control the type of cloud shown and whether precipitation occurs, SMILE runs all values are changed in the background. Even if the users do not interact with the sliders for a while, the clouds move with a wind speed that is not under the player's control. All sliders help users to create clouds, rain, and snow, but the complex part of the simulation runs without the user's control. SMILE is an interactive system that encourages an exploration of the virtual environment. From this point of view, SMILE can be considered a casual creator experience. Based on K. Compton's definition, "A Casual Creator is an interactive system that encourages the fast, confident, and pleasurable exploration of a possibility space, resulting in the creation or discovery of surprising new artifacts that bring feelings of pride, ownership, and creativity to the users that make them (Compton & Mateas, 2015)"

Aesthetics

Aesthetics is related to the overall environment of the game. It directly relates to the players' experience. To help players immerse in the game, SMILE includes unpredictability, interaction, and sounds.

The sounds include seagulls, ocean and waves, rain, and snow blizzard sounds (Below Figure). Sound has a significant effect on the mood and atmosphere of the game. Another sound is the sound of winning after creating all clouds and finishing tasks. This feature is based on one of the methods in Schell's lens of indirect control (Schell 2008). There are some other indirect controls in the game to help players interact on their own such as interface and visual design. The interface is transparent, and four sliders control the variables that influence the scenes. On indirect control, Schell asks designers "can I design my interface to get players to do this?" and "can I use visual design to get players to do it?" (Schell, 2008). The visual design and interface help players to easily interact with the game.

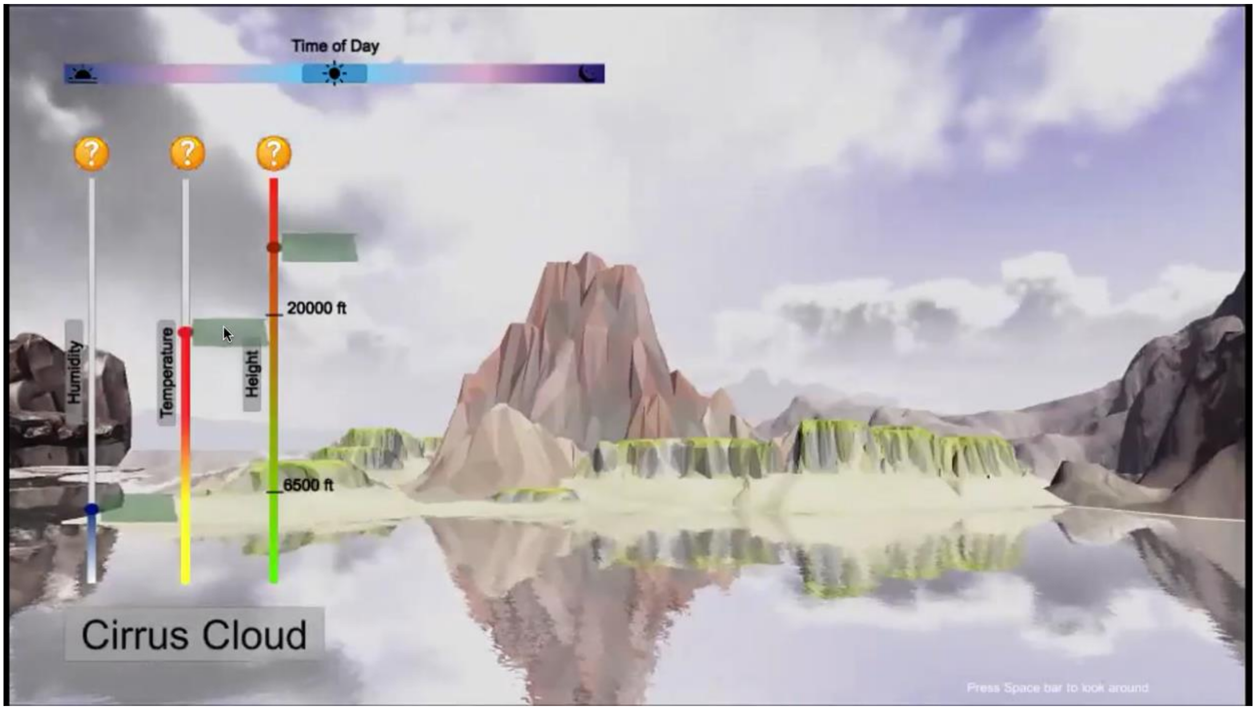


Figure 14:Sample interactivity, See https://youtu.be/fUum6_En1qM for a video

Another aesthetic element of an effective educational game is sensory stimuli. It refers to the interplay between graphics and sounds to appeal to players' senses. Referring to the above video, all sounds for different precipitations correspond to the respective visuals.

Pedagogy

Pedagogy differentiates educational games and entertaining games. Learning theory and pedagogical approaches are fundamental requirements to develop educational games. SMILE as an interactive educational experience combines good game design and pedagogy. The educational challenges in SMILE are according to the school curriculum. Therefore, it creates a good experimental flow (Flow theory). Flow questionnaire is a set of 10 questions measured by GEQ that is explained explicitly in the result chapter. All terminologies such as temperature, humidity, and name of clouds are familiar to players. In case they need more information, there are question mark buttons that provide a short description of sliders. Figure 5 shows an example for the temperature question mark. Also, based on the lens of the virtual interface (Schell, 2008), a game needs to deliver information to the players in a way that will not interfere with their interactions. That is why the information buttons (question mark buttons) are small, which are located on top of each slider, and players can open the information panel whenever they want. The panel is transparent, keeping the player immersed in the current scene, and does not interfere with any interactions in the game environment.

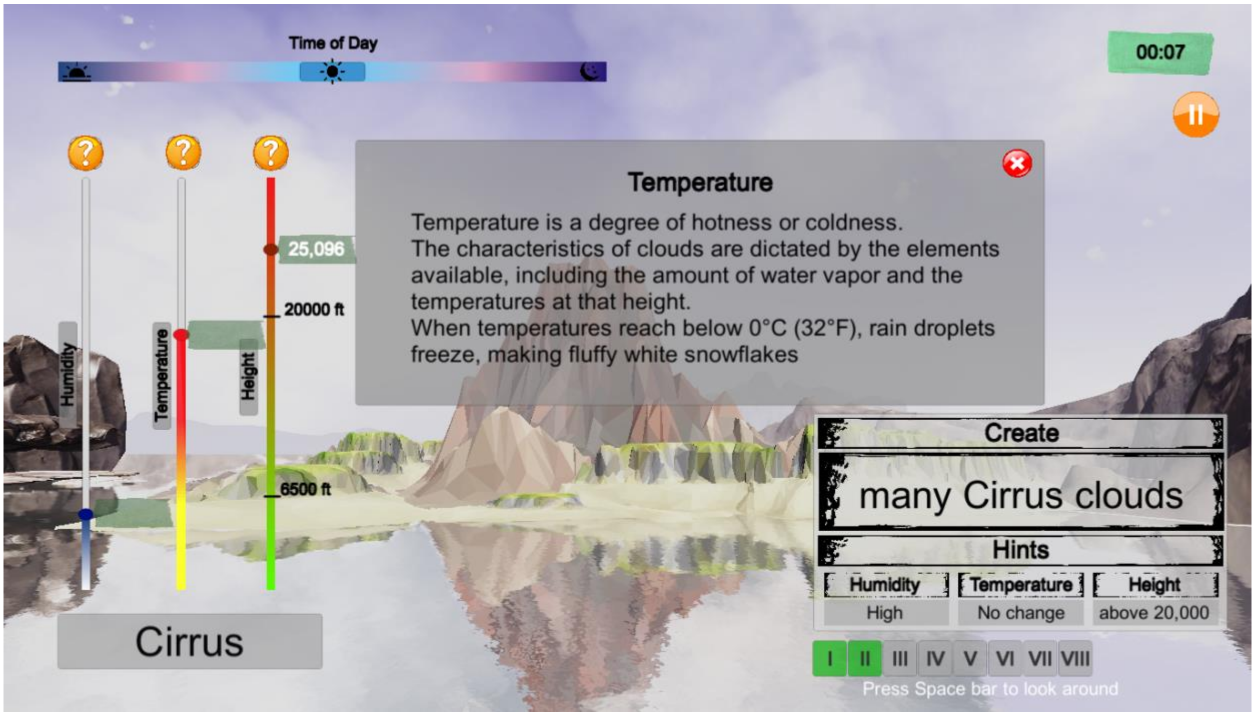


Figure 15: Example of question button to describe words

SMILE's players control the time of day, height, temperature, and humidity. Then the scene changes based on their decision. It generates intentional tension between the degree of control and freedom of interaction. This tension helps interactivity (Swartout & van Lent, 2003).

Gently guiding players with the create panel forms a mental framework for players to approach problem-solving and knowledge acquisition. The complexity level matches with players' age and general knowledge. The hint panel helps to maintain a balance between the game challenge and the players' abilities described as a comfort zone in flow theory.

Technology

One of the game design elements is technology. Technology's key component is adaption, which is a mechanical adjustment of the game. Based on Lope's definition (De Lope, Medina-Medina, Soldado, García, & Gutiérrez-Vela, 2017), adaption is the game's ability to adjust itself to the players' profile. Also, it relates to their method of interaction and educational challenge. The current version of SMILE does not have this adjustment; however, the design is matched with the players' performance.

This chapter presents a series of game design elements that the author applies in developing the educational simulation. SMILE is developed through a methodology starting with the identification of the weather science. See Appendix A and B for further details about the design and meteorological science. The next chapter details the thesis methodology and procedures. It includes the research setup and the instrument used to conduct the study, and method of analysis.

CHAPTER 4: METHODOLOGY

This chapter explores the thesis method, beginning with research questions, then continuing to the simulation design, data collection, and the instrument used in this study. The chapter ends up with explanations of analysis techniques.

Research Questions and Hypothesis

The research questions examine the relationship, if any, between the significant impacts of the designed simulation on student engagement as measured by GEQ and learning achievement. Correlation testing is employed to address two research questions and two hypotheses.

The following research questions are used to conduct this study:

Research question 1: Does SMILE produce any statistically significant impact on student engagement as measured by GEQ?

Research question 2: Is there any statistically significant impact by SMILE on student learning achievement?

Hypothesis 1: There is a statistically significant improvement in learning outcomes from pre-test to post-test.

Hypothesis 2: There is a positive correlation between learner engagement, as measured by GEQ, and the knowledge gains of a student in the topic after interacting with SMILE.

Simulation Design

In a classroom, teachers can often recognize the emotional state of their students and react accordingly. For example, if students are having some difficulty, the teacher can slow down the pace of that activity. Unlike in a classroom, in interactive digital learning environments, the learner’s emotions are not easily monitored. However, one way to guide exploration is using questions related to the experience's learning goals requiring a student more deeply engage with the game. Goal-oriented questions can also encourage curiosity and discovery. In SMILE, a create-hint panel was added to the game for this purpose. For instance, when the goal is to create a Cumulus cloud, the hint panel shows the values of humidity, temperature, and height needed.

SMILE starts with a simple statement and a short question (Figure 16). It then has two windows of tutorials before starting the interactive experience. It follows a design of gently guiding students through questions or statements presented on the screen (Adams et al., 2008; PhET, 2019). After the tutorial section, a panel on the bottom left side of the windows asks the player to create a different cloud every 30 seconds by following the hints. The final minute of the interactive experience asks students to explore anything in the simulation without explicit guidelines (Figure 17).

Figure 16: SMILE Welcome Screen

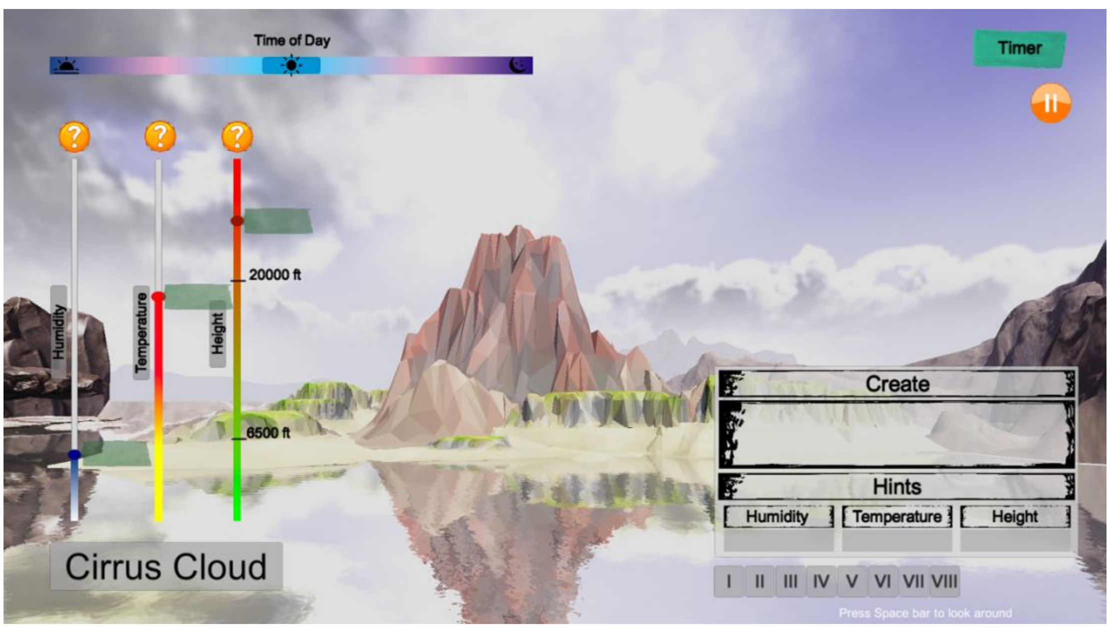


Figure 17: SMILE Main Window

Design is based on six properties of good digital games to promote learning defined by Gee (2009). The first focuses on the goal and problem-solving. The “create panel” (see Figure

11) in SMILE, asks the player to make a type of cloud. This study focuses on three main clouds: cumulus, altostratus, and cirrus. Their formation is based on the elevation of the clouds' base. The player can change the height using a vertical scrollbar in the game. Each cloud can be thicker or thinner based on the amount of humidity. Detail explanation is available in game design chapter. The second property, defined by Gee (2009), concerns how micro-control creates a feeling of influence. There are four sliders that players interact with: humidity, temperature, height, and time of day. All these controls show an immediate change in the simulation scene.

The third property Gee (2009) discusses is experimental learning opportunities. SMILE teaches learners to think strategically and see the immediate feedback during each step of the interaction. The fourth property is about flow theory and the importance of matching players' skills and knowledge. All cloud names and terminologies using in SMILE are based on elementary school textbooks. If the players do not remember the elevation of any clouds type or do not know that rain may change to snow at lower temperatures, the hints panel below the create panel helps them create what they are supposed to create. Therefore, there is a challenge to make different clouds, but it is matched with the players' knowledge. The fifth property is the need for the interaction experience to model a combination of the real-world and the abstract. Having a low-poly environment with realistic weather conditions and clouds satisfied this property. Finally, the sixth property from Gee (2009) is allowing the player to enact a different trajectory through the experience, thereby creating his or her own story. The 30 second time that shows as a countdown timer on top of the screen \increases the challenge; however, players are able to make their weather without following the create or hints panel. Moreover, they have one minute at the end of the test to explore more and create anything such as a rainbow or click on the spacebar and look around the simulation environment.

Study Design and Population

This study evaluated an educational interactive experience as a teaching tool for science for 4th to 6th grade students. It was approved by the Institutional Review Board (IRB) at University of Central Florida (UCF). The approval and closure letters can be found in APPENDIX G.

Test Setup

The test was conducted online, and a link to the game and all surveys and questionnaires was shared with participants and their parents. Each participant used their own device to play the game and fill out the surveys. All information, consent form, and meeting link were sent via email to the participant's parent or guardian.

All steps were explained by the researcher to each participant to guide them through the order of surveys (Figure 18). There was no interaction while students were answering questions or playing the game. There were no video recordings of any part of the procedure taken at any time. However, the researcher wrote down observations of the verbal and non-verbal reactions and behavior of participants. All information was coded, and no one except the researcher and PI of the project had access to the information.

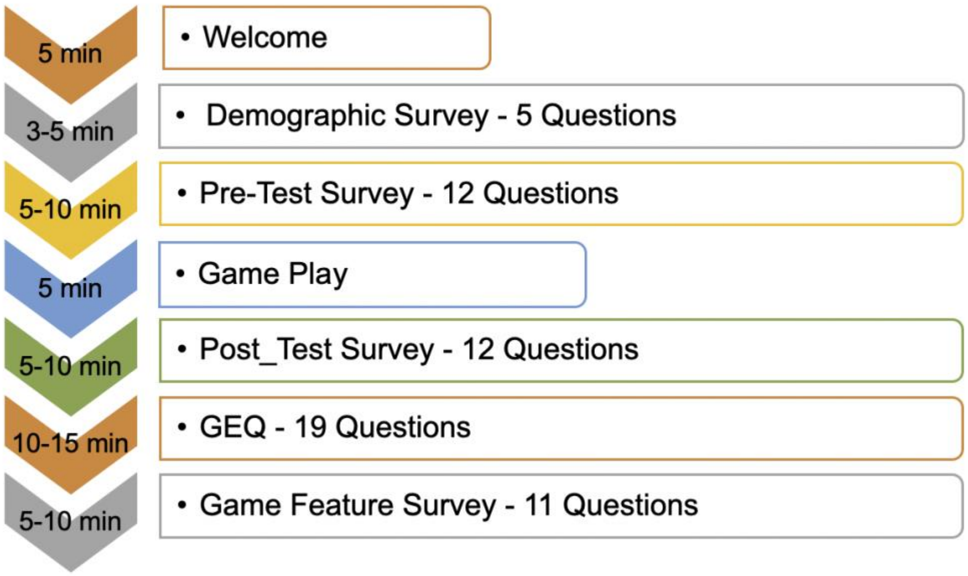


Figure 18: Study and Survey Ordering

Questionnaire and Response Measure

A self-reporting questionnaire with rating scales was used to collect data. Free discussion was not considered in the quantitative analysis. All the questionnaires used in this study were adapted to be understandable for children; the words and sentence structures were simplified. To achieve optimal responses, a Smiley Face Likert scale (Figure 19) was created based on existing studies (Hall et al., 2016; Stange et al., 2016). Although the number of players does not exceed 100 people and the left-side bias may not change the result, the author decided to display the options from negative to positive in a horizontal setup.

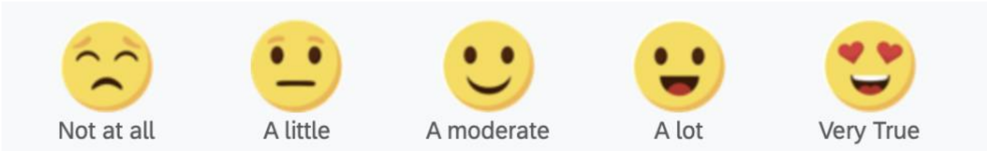


Figure 19: Smiley Face Likert

Game Engagement Questionnaire

Engagement has been identified as a critical element of learning in serious game research. One of the credible ways to measure engagement is a self-reported survey. In this study, the Game Engagement Questionnaire (GEQ) was used. The number of questions within the GEQ may vary in different studies; however, it always contains a core questionnaire that measures four elements of engagement: immersion, flow, presence, and absorption. All items in the GEQ are presented as statements and used a five-point Likert-style scale for participants to determine agreement or disagreement. The questionnaire consisted of 19 items in total. The Full GEQ can be found in APPENDIX C.

Demographic Questionnaire

The demographic questionnaire assessed some basic information about the participants. It aimed to understand the identity and characteristics of responders better. It included six questions about age, grade, gender, race, and previous experience with video game such as time spent playing. The full demographic questionnaire is in APPENDIX D.

Game Feature Questionnaire

The game feature questionnaire is a set of ten questions developed by the researcher of this project. This questionnaire is an evaluation of SMILE’s features. The questionnaire can be found in APPENDIX E.

Pre/Post-Test Questionnaire

The pre/post-test questionnaire measured the participant’s knowledge about clouds and the relationship between temperature and humidity in cloud formation and precipitation. The knowledge and questions were based on the Florida Statewide Science Assessment. For 5th grade, it was based on SC.5., “E.7.4 students distinguish among the various forms of precipitation and make connections to the weather in a particular place and time.” For 4th grade, it was based on SC.4., “P.8.2 students identify properties and common uses of water in each of its state and cloud types” (Florida Statewide Science Assessment, 2019). The pre/post-test questionnaires can be found in APPENDIX F.

Statistical Analysis

Two sets of statistical analysis were conducted on different aspects of this study. The first analysis was conducted to see if the students had any significant improvement in meteorological knowledge using this interactive experience and its interface. The data was based on a pre/post-test knowledge questionnaire. The second analysis examined the engagement level in the interactive experience based on GEQ. The IBM SPSS version 25 was used for statistical analysis. This study aimed to have a sample size of 30 elementary students. However, the crises of the COVID-19 virus pandemic limited the accessibility to the student. The final sample size was 23 students.

Statistical Analysis on Knowledge Test

The knowledge test consisted of 12 true or false questions about clouds and meteorological science. The correct answer had a score of 1, and the false answer had a score of 0. The minimum and maximum scores for each participant were 0 and 12, respectively. In terms of selecting a correct statistical test, the author considered the research hypothesis and the type of data. The pre- and post-data were skewed; the quantitative non-normal tests were chosen.

First, the author used the Wilcoxon Signed-Rank test to evaluate if there was a statistically significant improvement in a participant's average knowledge after interacting with SMILE. A detailed explanation is available in the next section.

Next, the Spearman test was used to specify if the two sets of knowledge tests had any significant correlation. This test showed if the initial knowledge of participants affected the improvement or not. In other words, the post-test and final knowledge increases regardless of what students knew before the test.

General Test

Before selecting the statistical test on pre- and post-test scores, the skewness of the data was estimated using the following formula.

$$Skew = \frac{n}{(n-1)(n-2)} \sum \left(\frac{x_i - \bar{x}}{s} \right)^3 \tag{1}$$

This test showed the skew for pre- and post-test were 1.4 and 0.42, respectively. Therefore, it is assumed at least the pretest is fully skewed, and the statistical analysis is carried out using nonparametric tests. Using Quartile 1, Quartile 2, and the following formula, the Upper Fence and Lower Fence for both pre- and post-data was estimated, and outlier data was specified. The results from 23 participants showed the pretest score for one is the outlier and that participant were removed from evaluations.

IQR = Quartile 2 – Quartile 1

Upper Fence = Quartile 2+(IQR x 1.5)

Lower Fence = Quartile 1+(IQR x 1.5)

Wilcoxon Signed Rank Test

The Wilcoxon Signed-Rank (WSR) test is a nonparametric statistical test to determine if there is a significant difference between two sets of related data without assuming the data is normally distributed (Wilcoxon, 1992). This test is appropriate since there are two sets of paired or repeated observations, which are skewed and not normally distributed. The author was interested in evaluating the effectiveness of the training provided by SMILE, and the Wilcoxon test analyzes the significance of the difference between the median of two sets of data (pre-test and post-test scores).

The null hypothesis (H0) for the Wilcoxon test indicates two samples have the same medians. The test starts with ranking the difference score for each participant. The ranks are assigned from 1 to n from the smallest to the largest absolute difference. The final step is based on the summation of rank numbers for each group.

Since the number of participants (n = 22) in this study is large enough (n(n+1)/2 > 20) for this test, a normal approximation is used with the following formula:

Here are the steps of the test when there are two sets of data, x1i and x2i, and the number of samples for each set is N:

1. Calculate the $|x_{2i} - x_{1i}|$ and $sgn(x_{2i} - x_{1i})$ where sgn is sign function:

$$sgn(x_{2i} - x_{1i}) = \begin{cases} -1 & if \quad (x_{2i} - x_{1i}) < 0 \\ 0 & if \quad (x_{2i} - x_{1i}) = 0 \\ 1 & if \quad (x_{2i} - x_{1i}) > 0 \end{cases}$$

2. Reduce pair size by removing the pairs where sgn is equal to 0 and the new sample size is named N_r
3. Rank the sets based on the absolute difference and use the average of rank for the tie values
4. Calculate the W and Z score (for the samples bigger than 20).

$$W = \sum_{i=1}^{N_r} [sgn(x_{2i} - x_{1i})];$$

$$\sigma_W = \sqrt{\frac{N_r(N_r+1)(2N_r+1)}{6}}; \qquad (2)$$

$$z = \frac{W}{\sigma_W}$$

The assumption for the sample sets were higher than 20 and the W followed a normal distribution with the above-mentioned z score and standard deviation σ_W. For the two-tailed test, the hypothesis was rejected when the critical z from the table was less than the absolute calculated z.

Spearman Rank Correlation Coefficient

The Spearman Rank Correlation Coefficient (SRCC) is a nonparametric method for measuring the correlation of ranks between two variables. SRCC varies between +1 and -1. The highest score is when there is a total positive linear correlation, 0 for no linear correlation, and the lowest score is when there is a total negative linear correlation.

The formula for SRCC is like the Pearson Correlation Coefficient with the only difference using the variable of two groups for ranking:

$$r_s = \frac{\sum_{i=1}^{n}(rgx_i - \overline{rgx})(rgy_i - \overline{rgy})}{\sqrt{\sum_{i=1}^{n}(rgx_i - \overline{rgx})^2}\sqrt{\sum_{i=1}^{n}(rgy_i - \overline{rgy})^2}} \qquad (3)$$

Where n is the sample size, rgx_i and rgy_i are the rank of individual sample points; $\overline{rgx}$ and $\overline{rgy}$ are sample rank mean for rgx_i and rgy_i.

The null hypothesis (H_0) and alternative hypothesis (H_1) of the significance test for correlation can be expressed in the following ways, depending on whether a one-tailed or two-tailed test is required.

Two-tailed significance test:

H_0: $r_s = 0$; the correlation coefficient is 0; there is no association.

H_1: $r_s \neq 0$; the correlation coefficient is not 0; a nonzero correlation could exist.

where r_s is the SRCC.

The sample size affects the significance of r. Figure 20 shows critical values of PCC for different sample size at the 0.05 level which can be used for SRCC (Soper, Young, Cave, Lee, & Pearson, 1917).

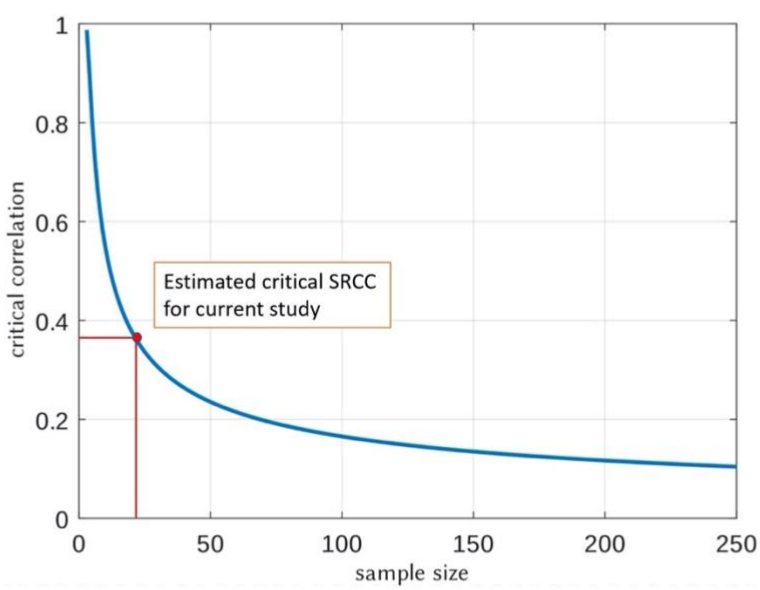


Figure 20: Critical Values of Pearson's Correlation Coefficient

Statistical Analysis on Engagement

Four different aspects of engagement (immersion, presence, flow, and absorption) are measured through 19 questions. The descriptive answers of "Not at all," "A little," "Moderate," "A lot," and "Very true" were evaluated based on the percentage of participants for each category. The answers from "Not at all" to "Very true" were converted to numbers 1 through 5, respectively. The Spearman correlation between the knowledge test and engagement was evaluated.

Hierarchical Cluster Classification

Hierarchical Clustering (HC) is a classification method that classifies the members based on their similarity. There are two major approaches for HC: top-down and bottom-up. In this study, the bottom-up (agglomerative clustering) approach was used. In this method, each observation assumes to be one group, and then in each step, a pair of clusters are merged to get one cluster on top of the hierarchy.

There are different methods to measure the distance between the two members. The Squared Euclidean Distance (SED) is one of the methods used in this study.

$$d = \sum_{i=1}^{n} (X_i - Y_i)^2 \tag{4}$$

The agglomerative clustering method has different branches for merging the clusters. The method used in this study was Ward's minimum variance or simply Ward's method. In this method, clustering is based on the minimum within-cluster variance. **Error! Reference source not found.** is used to minimize within-cluster.

$$ESS = \sum_{i} \sum_{j} \sum_{k} \left| X_{ijk} - \bar{X}_{i.k} \right|^2 \qquad (5)$$

Where ESS is Error of Sum of Squares, X_{ijk} is the value of variable k in observation j belonging to cluster I, and $\bar{X}_{i.k}$ is the average variable k in cluster i.

In this study, engagement questions (19 questions) were clustered based on the participant’s answers.

CHAPTER 5: RESULTS

After IRB approval, data collection was started and completed during summer 2020 entirely online via Zoom. Each experimental session was approximately 45 minutes, and 23 students participated. After the first statistical analysis, one outlier data point was found. The data is for a participant that got the pre and posttest 100% correct. The researcher excluded the outlier data and continued analysis with the data from 22 participants.

Study Participants

The population of this study was comprised of elementary students in Florida and specifically focused students going to 4th, 5th, and 6th-grade students. The sample included 30 students. After data evaluation, extraction of an outlier, and deleting missing data, the sample size was reduced to 22 students compromised of boys (n = 9) and girls (n = 13). The researcher used statistical analysis within SPSS version 25 for quantitative analysis and qualitative content analysis.

Analysis of Demographic Data

Descriptive statistics for the demographic data are summarized in Table 5. The ages varied from 8 to 12 years old. The participants were mostly girls (65%) and 52% of them had finished their 4th grade. The original number of participants was 23; one participant was disregarded in the statistical test due to being an outlier.

Table 5: Demographic Data of Study Participants

	n	%
Gender		
Female	15	59
Male	8	41
Grade		
Going to 4th grade	3	14
Going to 5th grade	12	55
Going to 6th grade	8	32
Age		
8-9 Years	2	9
9-10 Years	8	36
10-11 Years	5	23
11-12 Years	8	32
Time of play per day on touchscreen devices		
Less than an hour	2	5
1-2 hours	6	27
2-3 hours	7	32
More than 3 hours	8	36

Note. From 23 participants 1 of them is removed after outlier test

Race was not expected to play a role in this study. However, most of the participants (52%) preferred not to mention their race. The rest of the participants were white (35%), Asian (9%) and African American (4%). Since the skills of players was important, the play-hour per day for each player was collected as well. Based on the data, 5% of participants played less than an hour per day. Most participants played games on touchscreen devices for more than two hours a day.

Hypothesis Testing

The pre- and post-testing of meteorological knowledge resulted in two series of scores between 0 to 12. Each correct answer had a score of one and an incorrect answer was zero.

The Wilcoxon test was used to identify if there was a significant improvement from pretest to posttest. It showed if participants significantly improved their knowledge about clouds and precipitation by interacting with SMILE.

Error! Reference source not found. shows the average scores of 22 participants improved from 4.77 to 7.50. The standard deviation for the pretest was 1.798 and this value changed to 2.041 for the post-test. Figure 21 shows the whisker plot of pre- and post-test as a standardized way to display the distribution of the data.

Table 6: Descriptive Pre- and Post-test Statistics (n = 22)

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles 25th	Percentiles 50th (Median)	Percentiles 75th
Pretest	22	4.77	1.798	2	9	3.00	4.5	6.00
Posttest	22	7.5	2.041	4	11	6.00	7.00	9.00

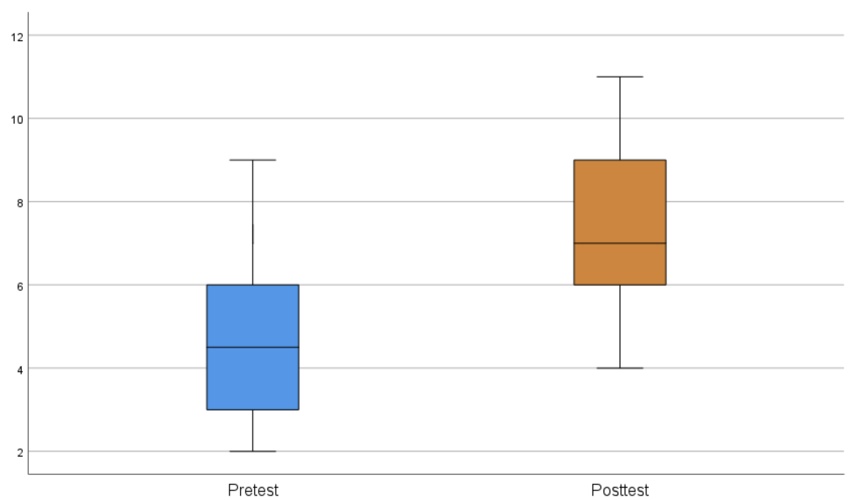


Figure 21: Pre- and Post-test Box-Whisker Plot (After Outlier Removal)

Table 7 and Table 8 show the pre- and post-test statistics and the results of the Wilcoxon test from SPSS. The rank table (Table 7) provides the output information of the Wilcoxon test. All positive ranks indicate post-test data has a higher score than pretest.

The Z value was -4.036 and the probability (p) was .000. This means the null hypothesis was rejected and there was significant development between pre- and post-test. The data indicated that post-test ranks were statistically significantly higher than pre-test ranks.

Table 7: Pre- and Post-test Ranks

Ranks

		N	Mean Rank	Sum of Ranks
Pretest-Posttest	Negative Ranks	0[a]	.00	.00
	Positive Ranks	21[b]	11.00	231.00
	Ties	1[c]		
	Total	22		

a. Posttest< Pretest
b. Posttest> Pretest
c. Posttest= Pretest

Table 8: Wilcoxon Test Result

	Posttest -Pretest
Z	-4.036[b]
Asymp. Sig. (2-tailed)	.000

a. Wilcoxon Signed Ranks Test
b. Based on negative ranks.

The second hypothesis suggested a positive relationship between the level of engagement as measured by GEQ and their knowledge achievement. The level of engagement for the interactive experience was evaluated using the simple graphs of the number of participants or percentage for each category of answers including “Not at all”, “A little”, “Moderate”, “A lot” and “Very True”. For each category of engagement (flow, presence, absorption, and immersion) different colors were used. Table 9 shows the results of the survey for each statement. Each aspect of engagement was specified using one to several statements. In this table, the mode and median are estimated using the equivalent scores of each category of answer ("not at all" is equal to 1, "very true" is 5, and the rest of the answers are between these two).

Table 9: GEQ Answer Percentages (n = 23), Each color represents one category; Blue: absorption, Orange: flow, Green: presence, Gray: immersion

GEQ	Not at all	A little	Moderate	A lot	Very True	Mode	Median
1. I lose track of time	22	13	39	9	17	Moderate	Moderate
2. Things seem to happen automatically	43	26	0	22	9	Not at all	A little
3. I feel different	22	30	35	0	13	Moderate	A little
4. I feel scared	87	13	0	0	0	Not at all	Not at all
5. The game feels real	0	9	22	39	30	A lot	A lot
6. If someone talks to me, I don't hear them	48	30	17	0	4	Not at all	A little
7. I get wound up	52	26	13	0	9	Not at all	Not at all
8. Time seems to kind of standstill or stop	30	22	30	9	9	Moderate	A little
9. I feel spaced out	30	39	9	13	9	A little	A little
10. I don't answer when someone talks to me	61	17	13	4	4	Not at all	Not at all
11. I can't tell that I'm getting tired	30	9	17	13	30	Very true	Moderate
12. Playing seems automatic	52	4	22	13	9	Not at all	Not at all
13. My thoughts go fast	4	43	13	22	17	A little	Moderate
14. I lose track of where I am	48	39	9	0	4	Not at all	A little
15. I play without thinking about how to play	17	26	22	22	13	A little	Moderate
16. Playing makes me feel calm	0	26	26	4	43	Very true	Moderate
17. I play longer than I meant to	17	22	30	9	22	Moderate	Moderate
18. I really get into the game	0	17	4	48	30	A lot	A lot
19. I feel like I just can't stop playing	13	9	26	35	17	A lot	A lot

A five-point Likert scale was used to indicate the disagreement or agreement on different levels. Based on this category, the first answer choice of “not at all” represented disagreement and the rest of the choices represent agreement with different significant level.

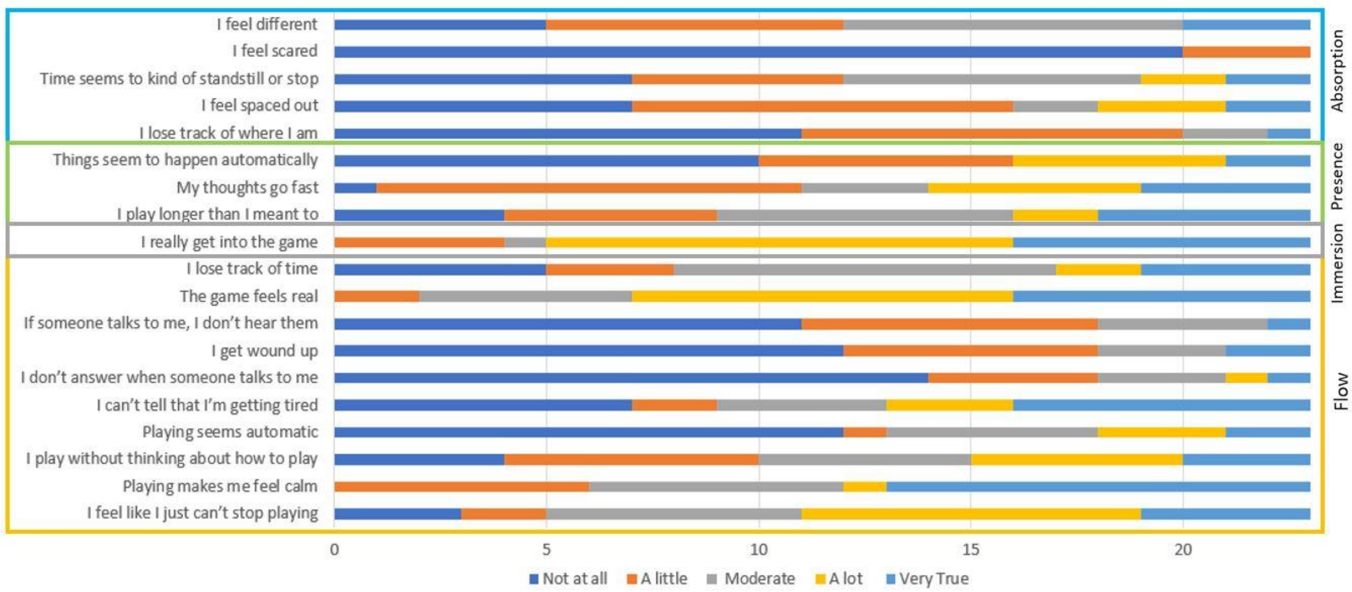


Figure 22: GEQ results (n=22)

Figure 22 shows the survey results for flow-engagement questions. Based on this figure and Table 9, six statements among 10 had partial to the full agreement (from “A little” to “Very true”). For the four statements of “If someone talks to me, I don’t hear them”, “I get wound up”, “I don’t answer when someone talks to me”, and “Playing seems automatic”, the common answer was no agreement or “Not at all”. It might be these questions were not designed for scientific games or participant were not able to understand them. For example, there is no reason a person by interacting with SMILE and knowing more about clouds would get wound up. The author believes the negative answers for some of these questions support engagement. The questions of “The game feels real” and “Playing makes me calm” showed the highest positive responses Generally, the results recognize how SMILE successfully helped the players to stay in the flow zone of engagement.

Above figure also shows the survey results for Presence-engagement questions. Based on these answers and Table 9, two of three questions had partial to a full agreement (from "A little" to "Very true"). For the question of "Things seem to happen automatically" the major answer was no agreement or "Not at all". Despite this being more difficult to answer for the scientific game, around 57% of participants were partially or fully agreed with engagement. These results show how SMILE successfully helped players to stay in the presence zone of engagement.

The results for absorption-engagement questions display a significant disagreement on questions of "I feel scared" and "I feel spaced out" from absorption. The explanation for these results is related to the genre of the learning interaction. It looks having this statement on GEQ to measure absorption on a calm educational interactive experience is not reasonable. It could also be participants felt emotionally safe because they participated from home and in the presence of their parents.

One statement in Figure 22 shows the survey results for immersion-engagement question. Based on this figure and Table 9, all participants agreed with the immersion statement. This indicates the participants felt immersion, one of the important factors of engagement.

Correlation Between Knowledge and Level of Engagement

The pre-, post-test, and difference of them were used to correlate data between engagement GEQ questions. For the GEQ, the answers are converted to numbers 1 to 5 for the answers of "Not at all", "A little", "Moderate", "A lot" and "Very true", respectively. Table 10 shows the results of the Spearman correlation between knowledge and GEQ scores. Based on this table, there is no significant correlation between pre-test and any of the questions. The post-test has a negative correlation with Question 2 (“Things seem to happen automatically”) and Question 15 (“I play without thinking how to play”) and a positive correlation with Question 3 (“I feel different”). The difference between pre- and post-test has a significant negative correlation with Question 15. This shows participants who disagreed with this question still showed improvement in knowledge of meteorological science.

Table 10: SRCC of Pre- and Post-test and Difference with Engagement GEQ

Correlations[c]

			Pretest	Posttest	Difference	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7	Q_8	Q_9	Q_10	Q_11	Q_12	Q_13	Q_14	Q_15	Q_16	Q_17	Q_18	Q_19
Spearman's rho	Pretest	Correlation Coefficient	--																					
		Sig. (2-tailed)	.																					
	Posttest	Correlation Coefficient	.462*	--																				
		Sig. (2-tailed)	.030	.																				
	Difference	Correlation Coefficient	-.345	.587**	--																			
		Sig. (2-tailed)	.115	.004	.																			
	Q_1	Correlation Coefficient	-.140	-.047	-.005	--																		
		Sig. (2-tailed)	.533	[illegible]	.981	.																		
	Q_2	Correlation Coefficient	-.227	-.428*	-.176	-.263	--																	
		Sig. (2-tailed)	.311	[illegible]	.433	.237	.																	
	Q_3	Correlation Coefficient	.206	.432*	.182	.268	-.208	--																
		Sig. (2-tailed)	.358	[illegible]	.418	.228	.352	.																
	Q_4	Correlation Coefficient	.170	.074	.053	-.086	.133	-.163	--															
		Sig. (2-tailed)	.449	.743	.814	.702	.554	.468	.															
	Q_5	Correlation Coefficient	-.145	.218	.346	.071	-.261	.273	.022	--														
		Sig. (2-tailed)	.519	.330	.115	.755	.240	.219	.923	.														
	Q_6	Correlation Coefficient	-.244	-.004	.143	.568**	-.255	.287	-.181	.555**	--													
		Sig. (2-tailed)	.273	.986	.525	.006	.252	.195	.420	.007	.													
	Q_7	Correlation Coefficient	-.030	-.086	-.140	.489*	.057	.135	.170	.024	.224	--												
		Sig. (2-tailed)	.895	.705	.534	.021	.800	.549	.451	.914	.316	.												
	Q_8	Correlation Coefficient	.019	.049	.112	.375	-.075	.462*	.368	.207	.262	.521*	--											
		Sig. (2-tailed)	.932	.829	.619	.085	.738	.030	.092	.355	.240	.013	.											
	Q_9	Correlation Coefficient	.110	.230	.090	.330	.063	.323	.000	-.003	.205	.635**	.515*	--										
		Sig. (2-tailed)	.628	.303	.690	.134	.781	.143	1.000	.989	.361	.001	.014	.										
	Q_10	Correlation Coefficient	-.147	-.129	.025	.661**	-.166	.275	.012	.168	.690**	.565**	.585**	.494*	--									
		Sig. (2-tailed)	.513	.568	.911	.001	.462	.216	.959	.454	.000	.006	.004	.020	.									
	Q_11	Correlation Coefficient	.269	.347	.035	.266	.058	.474*	-.054	-.070	.315	.284	.284	.737**	.413	--								
		Sig. (2-tailed)	.226	.114	.877	.231	.798	.026	.811	.758	.153	.201	.201	.000	.056	.								
	Q_12	Correlation Coefficient	.088	-.253	-.381	-.304	.288	.118	-.344	-.047	.009	-.063	-.273	-.061	-.046	-.057	--							
		Sig. (2-tailed)	.697	.255	.080	.169	.194	.601	.117	.836	.968	.780	.218	.786	.837	.803	.							
	Q_13	Correlation Coefficient	.311	.076	-.230	-.275	.036	.280	.121	.058	.086	-.033	.180	.023	.120	.264	.371	--						
		Sig. (2-tailed)	.159	.737	.304	.216	.875	.207	.590	.797	.705	.885	.424	.918	.595	.235	.089	.						
	Q_14	Correlation Coefficient	.125	-.068	-.172	-.074	-.101	.181	.182	.136	.170	.304	.250	-.029	.322	.030	.405	.586**	--					
		Sig. (2-tailed)	.579	[illegible]	.444	.743	.654	.419	.417	.546	.449	.168	.262	.898	.144	.896	.062	.004	.					
	Q_15	Correlation Coefficient	.004	-.453*	-.464*	-.015	.379	-.175	.064	.020	.208	.142	-.071	.048	.233	.214	.236	.533*	.220	--				
		Sig. (2-tailed)	.985	.034	.029	.947	.082	.437	.777	.931	.354	.528	.753	.831	.298	.339	.290	.011	.325	.				
	Q_16	Correlation Coefficient	.132	.266	.176	.015	.178	.244	.078	.032	.076	-.234	.191	.080	-.123	.157	.024	.236	-.060	-.044	--			
		Sig. (2-tailed)	.559	.231	.434	.946	.428	.274	.730	.888	.736	.295	.394	.722	.584	.486	.915	.291	.790	.846	.			
	Q_17	Correlation Coefficient	.354	-.074	-.374	.378	.060	.086	.471*	.102	.274	.208	.359	-.085	.257	.121	-.078	.315	.249	.267	.177	--		
		Sig. (2-tailed)	.106	.744	.087	.083	.791	.702	.027	.652	.217	.352	.101	.708	.249	.593	.730	.154	.264	.231	.431	.		
	Q_18	Correlation Coefficient	.248	.388	.153	.166	-.414	.434*	-.056	.211	.209	-.301	.121	.074	.138	.250	.011	.372	.192	-.024	.568**	.029	--	
		Sig. (2-tailed)	.265	.074	.495	.462	.055	.044	.805	.347	.349	.174	.590	.743	.542	.262	.961	.089	.392	.914	.006	.897	.	
	Q_19	Correlation Coefficient	.133	.146	-.004	.161	-.197	.375	-.032	-.022	.144	-.066	.326	.061	.223	.114	-.021	.479*	.268	.012	.656**	.126	.717**	--
		Sig. (2-tailed)	.554	.515	.984	.475	.378	.086	.886	.923	.524	.770	.138	.787	.319	.614	.924	.024	.228	.959	.001	.575	.000	.

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

c. Listwise N = 22

The SRCC was evaluated using the total scores of each category of engagement (Immersion, presence, flow, and absorption). Table 11 represents the SRCC for pre-test, post-test, and difference of immersion, presence, flow, and absorption values. Based on this table, there is no significant linear correlation between pre-test, post-test, or differences between categories of engagement. However, there is some positive correlation between these categories.

Table 11: Correlation of Immersion, Presence, Flow and Absorption Values

Correlations

			Pretest	Posttest	Difference	Immersion	Presence	Flow	Absorption
Spearman's rho	Pretest	Correlation Coefficient	--						
		Sig. (2-tailed)	.						
		N	22						
	Posttest	Correlation Coefficient	.462*	--					
		Sig. (2-tailed)	.030	.					
		N	22	22					
	Difference	Correlation Coefficient	-.345	.587**	--				
		Sig. (2-tailed)	.115	.004	.				
		N	22	22	22				
	Immersion	Correlation Coefficient	-.140	-.047	-.005	--			
		Sig. (2-tailed)	.533	.836	.981	.			
		N	22	22	22	22			
	Presence	Correlation Coefficient	-.100	.164	.297	-.005	--		
		Sig. (2-tailed)	.658	.467	.179	.981	.		
		N	22	22	22	22	22		
	Flow	Correlation Coefficient	.151	.109	-.036	.311	.419	-	
		Sig. (2-tailed)	.503	.630	.872	.159	.052	.	
		N	22	22	22	22	22	22	
	Absorption	Correlation Coefficient	.265	-.139	-.373	.060	.154	.291	--
		Sig. (2-tailed)	.232	.537	.087	.790	.493	.190	.
		N	22	22	22	22	22	22	22

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Game Features and Content Analysis

Recorded observations and the game feature questionnaire were used for descriptive data. After the post-test, participants answered the game feature questionnaire, which was designed by the researcher and based on related articles on game design. It used a five-point Likert scale with a minimum of 1 and a maximum of 5. The participants indicated how they felt while playing the game (Figure 23). These results indicated the entire experience was pleasant and satisfied the user. However, not all participants found the questions useful. All observations were conducted online via Zoom. The researcher wrote down all reactions. No audio or video took place during the sessions.

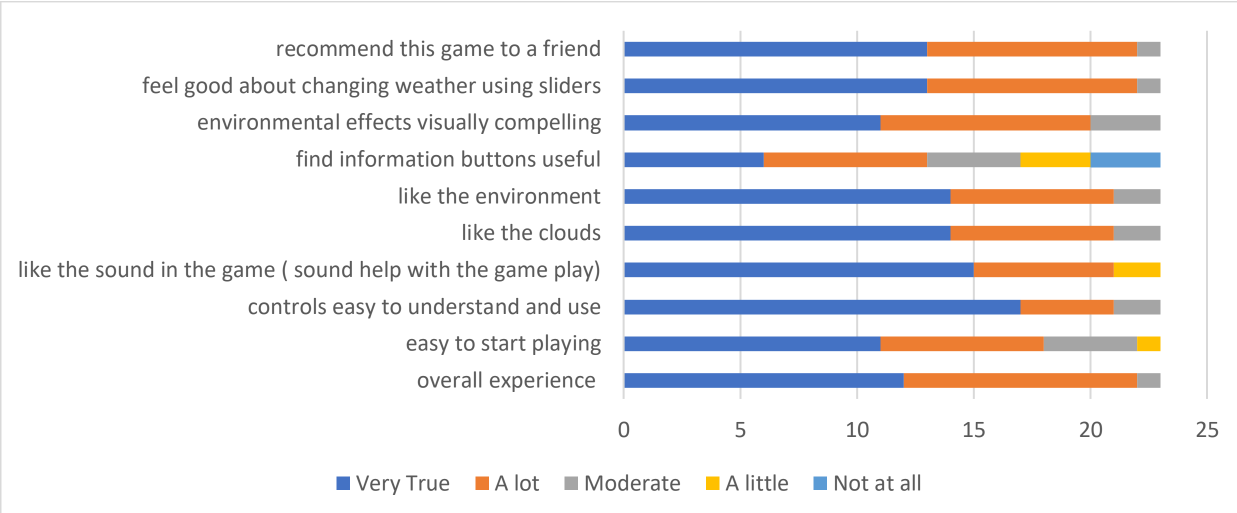


Figure 23: Game Feature Survey Result (n = 23)

Interaction Time with SMILE

SMILE contains eight, 30-second clues to make different clouds. At the end of this four-minute interactive experience, participants have one more minute to explore whatever they want. When they see the “Well Done” message with a button to take the survey, the game is over (Figure 24). 12 out of 23 participants continued interacting with SMILE without paying attention to the survey button.

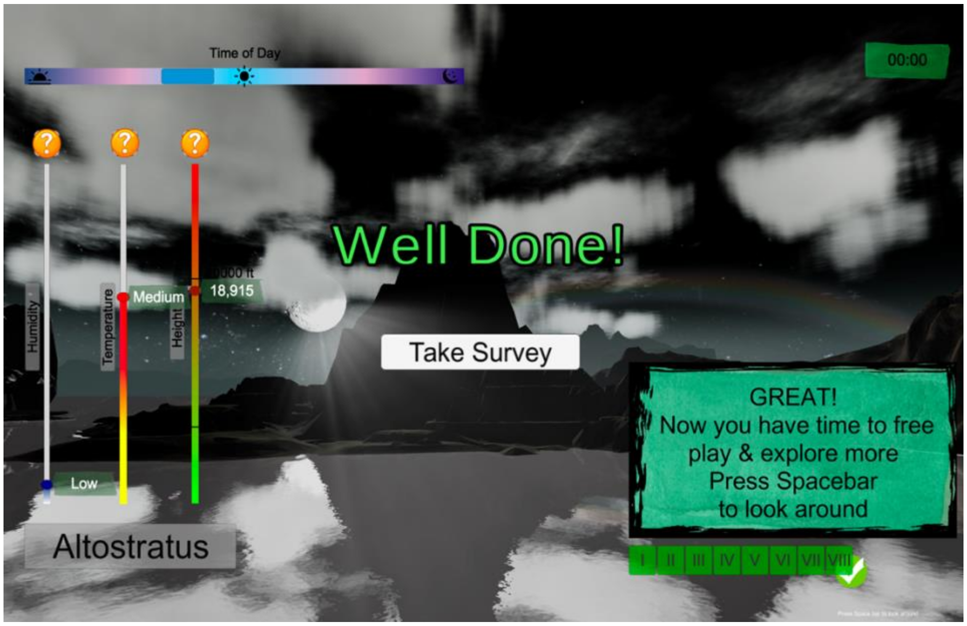


Figure 24: Sample Scene of Game Ending

Participants can be divided into three groups in how they handled the ending. Five participants did nothing or only clicked up to five times after the game ended. Six participants played for the last minute and then clicked on survey quickly. 12 (of 23) participants continued playing even after the messages appeared.

The most appealing scenes based on participants verbal and emotional reactions are the snowy scene and the rainbow making scene (Figure 25).

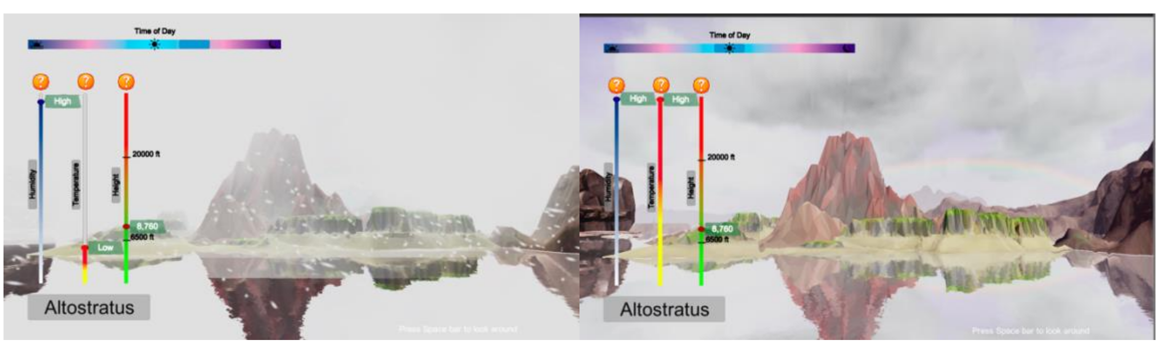


Figure 25: SMILE Sample Scenes of Snow and Rainbow

Discussion of these analysis is explored in the next chapter. Each hypothesis is discussed based on the result, followed by clustering method on the data. Recommendations and suggestions for the educators are provided to offer ideas on design the educational simulations and measuring the engagement.

CHAPTER 6: DISCUSSIONS

One of the main purposes of measuring game engagement is to evaluate its strengths and weaknesses. This study measured game engagement and asked about the game features to reveal any relationship between game design and engagement.

This study examined the importance of game design in engaging students in learning science. The main two questions that shape future research on learning interactive experience are the relationship between engagement and learning and the importance of designing a learning experience.

Designing SMILE

In this study, SMILE was designed and built using the Unity game engine by the researcher. The game feature survey results revealed how the game demonstrates the main properties of a good digital game for promoting learning (Gee, 2009). SMILE is based on real-world scenarios, and the clouds look realistic. There is no score in this experience, and all participants can feel they were a winner at the end of the experience. This matches how McGonigal (2011) describes how games focusing on real events help students to explore more than accomplishing the goal of the game. Based on the feature survey result, all participants found the clouds and environment visually compelling. Additionally, they felt they could control the weather in the game environment. This connects to the importance of micro-control in game design as discussed by Clark (1997).

As explained in the literature review chapter, the inverted-U model of pressure and performance applies to game design (Schell, 2008). To make the learning efficient, a level of guidance was added to SMILE in the form of the create panel (Moreno & Mayer, 2005). Using this panel, a player can follow the hints to create a new type of cloud every 30 seconds. Adding a timer to make each cloud helped the user not only have a structure to follow, but also put enough pressure on them to interact with SMILE. Without gently guiding the player, these statements would not be engaging.

There are three question marks above humidity, temperature, and height scrollbars in SMILE. By clicking on them, a pop-up screen shows the definition of each term and its effect on precipitation. About 40% of the participants did not notice the presence of question bars. Based on the researcher observation, only two participants clicked on the question marks (Figure 26).

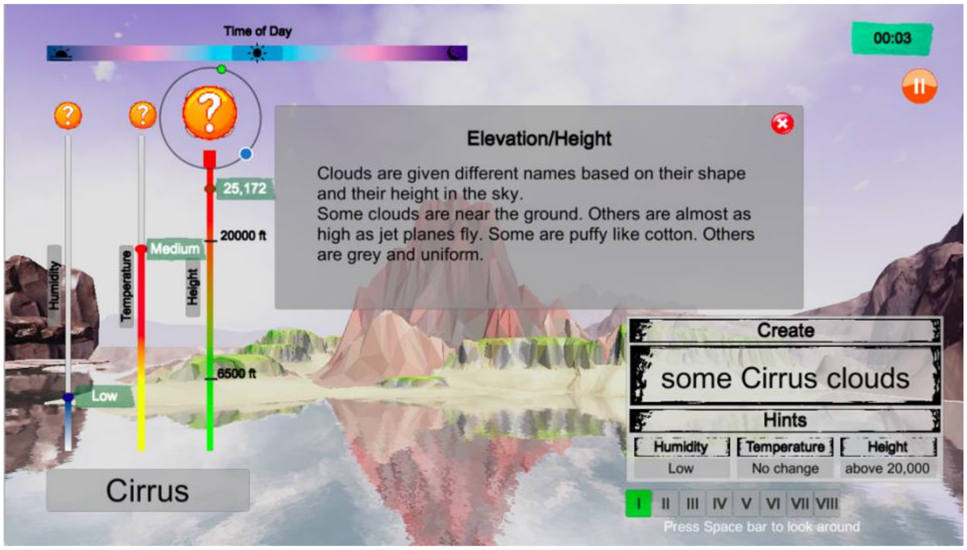


Figure 26: SMILE Sample Screen Scene

Digital interactive environments promote engagement and knowledge acquisition in the classroom (Emmanuel Tsekleves, 2016; Lameras et al., 2017a). The information presented in the game and the knowledge questionnaire are all based on 4th and 5th Sunshine State Standards science textbooks and can be easily applied to the classroom. Prior games and interactive experiences about weather forecasting mostly used two main meteorological diagrams and formulas. The first option worked with Skew-T, which are used by meteorologists and can be difficult for younger learners to understand. The others focused on factors in precipitation, temperature, and humidity. SMILE combined the effects of elevation on clouds and how temperature and humidity interact to make precipitation.

Hypothesis 1: Learning Outcomes

Participants showed significant improvement from pre-test to post-test for learning outcomes. This shows participants learned about the science of clouds and weather after interacting with SMILE. This result also supports the first hypothesis of how SMILE can have positive effects for learning outcomes (Figure 27). The result is clear and is supported by WSR statistical test as well.

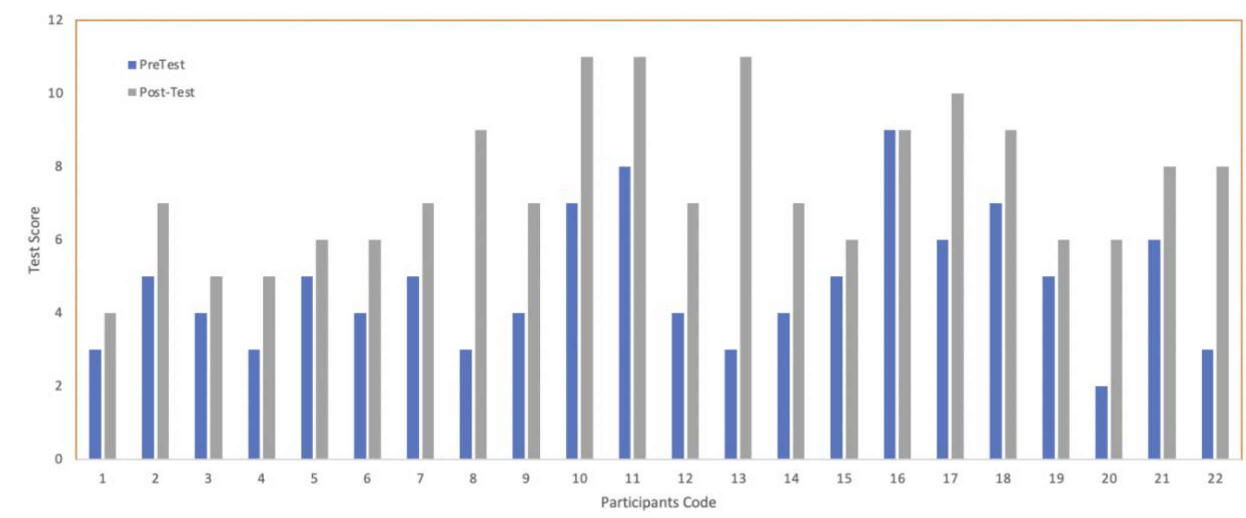


Figure 27: Comparison of Pre- and Post-test Scores (n = 22)

Hypothesis 2: Engagement

Previous research used GEQ to measure engagement in different types of games. Brockmyer et al. (2009) made the original questionnaire measuring engagement in violent video games. Since this original GEQ was designed for violent video games, some of the questions were not appropriate for scientific interaction games. For example, participants should not be wound up or scared when playing SMILE. Other researchers analyzing engagement in scientific games have followed this pattern of only using some of the GEQ or variations of the questions for their own studies (M. Chang, Evans, Kim, Deater-Deckard, & Norton, 2014; Zulehmay et al., 2019). An analysis of the responses to the GEQ questions which did not confuse participants suggests most of the GEQ participants had a good level of engagement. The test results and verbal behavior of the participants also showed not all the students were able to understand the statements on the GEQ. The most controversial responses were “I get wound up”, “I feel spaced out”, “my thoughts go fast”, and other negative statements. Half of the participants asked about the meaning of the mentioned statements.

We also evaluated the summation of engagement for each category by assuming an answer of “not at all” is full disagreement and the rest of the answer were partial or full agreement. Figure 28 represents this evaluation with about 70 percent of the answers supporting partial to full agreement on engagement. This is the total number of answers for each category, including the statements that were not proper for SMILE like “I feel scared.”

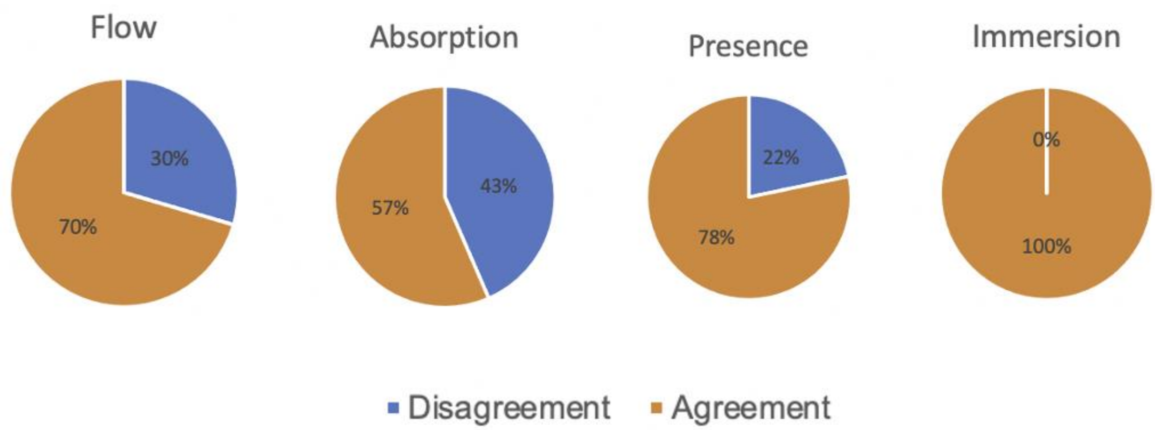


Figure 28: Percentage of Agreement vs. Disagreement for Engagement

Clustering

Generally, each GEQ statement should show different aspects of the engagement. Since there was a correlation between different GEQ and different categories of engagement (Table 10 and Table 11), the author tried to classify the engagement questions using the question scores and Hierarchical cluster analysis (Figure 29). This classification resulted in four clusters based on within-cluster minimum variance and maximum variance between clusters (Ward's method). Using these clusters, the SRCC of the summation of question scores for each cluster and pre-test, post-test, and the difference were evaluated. The results are presented in Table 12. This result shows the difference between pre-test and post-test had a significant negative correlation with cluster 2, which is the summation of scores for Questions 2, 12, 13, and 15. The questions mentioned in cluster 2 are all asking if participant feels everything is going so fast and automatically. That is why participant's answers are similar and cluster analysis puts all of them in one category. Since this simulation is educational and most of the participants didn't think everything was going automatically and fast like action games, so the answers were negative. The negative correlation between difference (posttest minus pretest) and cluster 2 shows that whoever thought the game is not automatic and they control the weather environment had better improvement.

There was no correlation between the difference and the original category of engagement (flow, presence, absorption, and immersion).

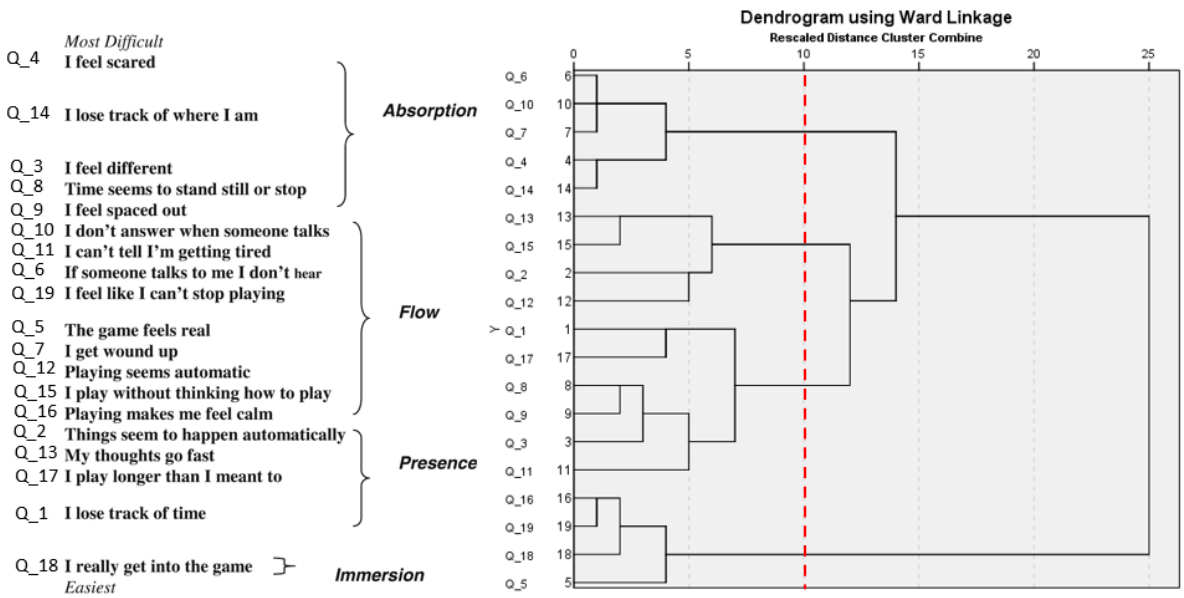


Figure 29: Results of Hierarchical Cluster Analysis on Engagement Question Scores

Table 12: Correlation of Pre-test, Post-test, and Difference with Clusters

Correlations[c]

			Pretest	Posttest	Difference	Cluster_1	Cluster_2	Cluster_3
Spearman's rho	Posttest	Correlation Coefficient	.462*					
		Sig. (2-tailed)	.030					
	Difference	Correlation Coefficient	-.345	.587**				
		Sig. (2-tailed)	.115	.004				
	Cluster_1	Correlation Coefficient	.051	-.034	-.095			
		Sig. (2-tailed)	.821	.881	.674			
	Cluster_2	Correlation Coefficient	.060	-.381	-.447*	.147		
		Sig. (2-tailed)	.791	.080	.037	.513		
	Cluster_3	Correlation Coefficient	.207	.168	-.040	.640**	.061	
		Sig. (2-tailed)	.356	.455	.858	.001	.787	
	Cluster_4	Correlation Coefficient	.109	.317	.181	.160	.070	.261
		Sig. (2-tailed)	.629	.150	.419	.477	.758	.240

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

c. Listwise N = 22

Confidentiality and Data Security

All data was collected through UCF Qualtrics and exported to SPSS for analysis. Data files were password protected and stored on UCF OneDrive and only accessible by the researcher and the PI. Personally-identifiable information was not collected in this study and demographic data was anonymous. All participants had a separate code created by the researcher for analysis. Each testing session was supervised and monitored by the researcher, but no video or audio recordings were created.

Conclusions

Educational games are a powerful learning tool. Students can enjoy playing the game while also learning some concepts that may not be interesting through traditional education methods. This study confirmed significant improvement in the knowledge of meteorology and clouds by students playing with the interactive educational simulation (SMILE), which was developed for this purpose.

While students were satisfied with game features, only 70 percent agreed they were engaged. The GEQ test scores showed that not all of these questions are appropriate for a scientific game like SMILE. Based on reactions and feedback of participants, some of the questions were not designed for the target age of this study (8 to 12 years old). Some statements such as “I feel scared” do not necessarily apply to educational games. Indefinitely phrased items such as “I feel different” did not have strong correlation or clear meaning to some respondents. Future studies should be selective or customize the GEQ questions to better evaluate engagement for different types of games. The absolute negative result of the statement “things happen automatically” indicates the degree of choice and control. SMILE provides players a sense of agency. Interactivity helps player to affect and change the simulation world.

Finally, for the first time, hierarchical clustering was used to classify the GEQ developed by Brockmyer et al. (2009). This classification was based on the test scores and tried to put the questions with the highest similarity in one group and questions with maximum variance in different clusters. The output was four clusters, with one of them showing a negative correlation with student knowledge improvement.

The current version of SMILE is based on the elementary school textbooks. A future version of SMILE could include more information SMILE that has more information and more options to control or change the weather. For example, changing latitude affects the temperature and the design will be set on the temperature limitations on each location on Earth. Moreover, the humidity and temperature range on mountains are different from the coastal area that can be applied to the next version of SMILE.

Ultimately, SMILE was an engaging and effective educational simulation. Future work could explore the potential for applying the techniques demonstrated here, both in design and evaluation, in other domains and for other age groups. This thesis presented a path and demonstration that will pave the way for those projects

APPENDIX A: CREATE AND HINT PANEL

Table 13: SMILE's create and hint panel

Type of cloud	Humidity	Temperature	Height
Some Cirrus clous	No change	No change	Above 20000ft
Many Cirrus clouds	High	No change	Above 20000ft
Altostratus	No change or Low	No change or High	Between 6500 and 20000ft
Rainy Altostratus	High	Medium	Between 6500 and 20000ft
Snowy Altostratus	High	Low	Between 6500 and 20000ft
Cumulus	No change or Low	No change or High	Below 6500
Rainy Cumulus	High	Low	Below 6500

APPENDIX B: SCIENCE FACTS IN SMILE

Interactive scrollbars

There are four scrollbars for time of day, humidity, temperature, and height. Except time of day slider that is horizontal, the same as sun movement, the rest of the sliders are vertical. Time of day changes from 00:00 am to 12:00 pm. The humidity and temperature sliders go from low to medium and high. The height goes from zero to 30000ft. The height's slider divides to below and above 6500ft and 20000ft to show different types of clouds.

Elevation/Altitude/ Height

The main factor in cloud categories is elevation. Clouds changes according to their height from the ground. If the user chooses lower values of the slider, the outcome will be a low-level cloud. The low-level clouds are stratus, cumulonimbus, cumulus, and stratocumulus. For the first version of SMILE, cumulus is the only cloud shown in low altitudes. If the user chooses heights between 6500 to 20000 ft. the cloud will be altostratus as a middle-level cloud. The third category of cloud which is elevated above 20000ft is cirrus. There are many other kinds of clouds in different atmospheric levels, but this project focused on three main clouds that are easy to find in the sky.

Humidity

Clouds are made of ice crystals (solid particles) or droplets (liquid particles). The potential of cloud formation depends on the amount of water vapor in a parcel of air named humidity. Temperature determines how much water the air can hold. The higher temperature can hold more water.

Relative Humidity

How much you feel the weather is wet or dry? The answer is the amount of relative humidity that is expressed as a percent. It is absolute humidity relative to the maximum humidity that is dependent on the temperature.

Example with numbers

A 100% relative humidity means the air is saturated. And above this level at the same temperature, clouds will appear. With unchanging humidity and increasing temperature, the relative humidity will decrease, and the cloud will be disappeared. For example, the air with 22°C (71.6°F) and 55% relative humidity would have 9g/Kg absolute humidity. It means that there are 9 grams of water per kilogram of air. Higher temperature leads to more humidity. So, 30°C (86°F) in temperature with the same absolute humidity will have 34% RH (Relative Humidity). But lower temperatures such as 12°C (53.6° F) with the same absolute humidity will have 100% RH, which means a cloud is formed (What is Relative Humidity Absolute Humidity and Dew Point, 2014).

If there is 100% humidity (Same RH) in the atmosphere, low temperature squeezes the water, and precipitation will occur. If the temperature is lower than the freezing point (0°C, 32° F), precipitation will be snow. Otherwise, it is rain. Following chart shows the relation between temperature and water vapor in different RH:

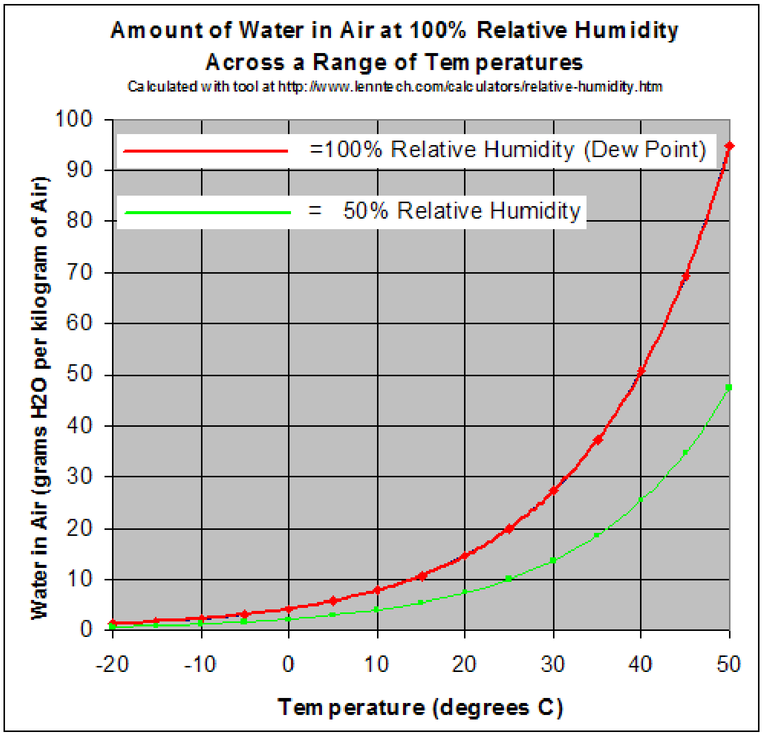


Figure 30: Relative humidity and dew points across a range of scales, by Greg Benson with calculations done on the site http://www.lenntech.com/calculators/relative-humidity.htm

For this interaction experience design, relative humidity keeps constant (100%). The user can change the temperature to get the saturated value or change the absolute humidity to start making the cloud. While the cloud formed, the decreasing temperature will cause precipitation. Making the cloud is eliminated from the current version of SMILE to make it easier for younger players. Clouds are already in the game environment, and players can change their types.

Math

This part explains the formula to form a cloud and occur precipitation. As described above, the current version already has a cloudy sky, and the formula is used to make the environment rainy or snowy.

Saturated water vapor: Ps in Pascal

Actual water vapor: Pa in Pascal

Temperature: T in Celsius

Relative humidity: RH in percentage

RH is the ratio of actual water vapor pressure to the saturation water vapor pressure at parcel temperature.

$$RH\% = {}^{P_a}/_{P_s} \times 100$$

The empirical equation given below is a good approximation of saturated water vapor at a specific temperature within the limits of the Earth's climate:

$$P_s = 610.78 \exp({}^{T}/_{(T\,+\,238.3)} \times 17.2694)$$

Based on this formula, every time that user changes T, Ps will be calculated by script, with looking at Pa (that it is a changeable value by user), RH will be known. If RH is less than 100%, there is no cloud in the scene. If RH is more than 100%, the cloud will appear in the scene.

Continue changing humidity to a lower value may make the cloud disappear. Continue changing humidity to a higher value will lead to having a denser cloud. Continue changing the temperature to a lower value makes clouds have rain or snow. Precipitation will be rain if the temperature is less than 0°C (32F); otherwise, it is rain. Continue changing the temperature to a higher value will make the cloud to be disappeared.

<u>Time of the day</u>

This factor is added as an independent element in cloud formation. Knowing that the time of the day typically affects humidity and temperature, in this project, it doesn't have any effect on other factors. In the real world, the temperature usually drops, and relative humidity goes up during the night. But these are not considered in this project. Day and night cycles will be used to change the time of day that affects clouds' color.

Clouds

Unlike the scene that is low poly, clouds are volumetrics with realistic shapes and shaders. The clouds are chosen from the Unity Asset Store. It is sky master ultimate by Nasos Tsatsos from the Artengame team.

In conclusion, many factors affect clouds formation and precipitation, such as time of the year, geographical location, and cloud condensation nuclei. This project aims to reduce these factors into three main elements, so it is understandable for younger learners.

APPENDIX C: GEQ

Table s: Game Engagement Questionnaire adapted from Brockmyer et al., 2009. Responses are measured on a five-point Likert scale with a corresponding to not at all and 5 corresponding to very true

Please indicate how you felt while playing the game:
I lose track of time
Things seem to happen automatically
I feel different
I feel scared
The game feels real
If someone talks to me, I don’t hear them
I get wound up
Time seems to kind of standstill or stop
I feel spaced out
I don’t answer when someone talks to me
I can’t tell that I’m getting tired
Playing seems automatic
My thoughts go fast
I lose track of where I am
I play without thinking about how to play
Playing makes me feel calm
I play longer than I meant to
I really get into the game
I feel like I just can’t stop playing

APPENDIX D: DEMOGRAPHIC SURVEY

Thank you for joining our research. Please answer these questions:

Q1. What is your age?

8-9 Years

9-10 Years

10-11 Years

11-12 Years

Q2. Which grade are you going to?

I am going to 4th grade

I am going to 5th grade

I am going to 6th grade

Q3. What is your gender?

I am a boy

I am a girl

I’d rather not say

Q4. What is your race?

I am:

White

Black or African American

American Indian or Alaska Native

Asian

Native Hawaiian or Pacific Islander

Other OR prefer not to say

Q5. How many hours <u>per day</u> do you play on touchscreen devices (Mobile phone, iPad, Tablet? ...)

Less than an hour

1 hour - 2 hours

2 hours - 3 hours

More than 3 hours

Q6. I am going to play this current game on a -----

Touchscreen Laptop

Tablet

iPad

Mobile phone

Laptop or PC with mouse

None of the above, please mention what is your device

____

APPENDIX E: GAME FEATURE SURVEY

Table 14: Game feature questionnaire, developed by author

Please indicate how you felt while playing the game:
How would you rate your overall experience in the game?
Did you find it easy to start playing the game?
Were the controls easy to understand and use?
How much did you like the sound in the game (did the sound help with the game play)
How much did you like the clouds?
How much did you like the environment in the game?
How much did you find information buttons (Questions marks in the game) useful?
Were the environmental effects visually compelling for the experience?
When you moved the sliders, you could tell they were changing things in the environment, did this feel good to you?
Would you recommend this game to a friend?

APPENDIX F: PRE-TEST/POST-TEST QUESTIONNAIRE

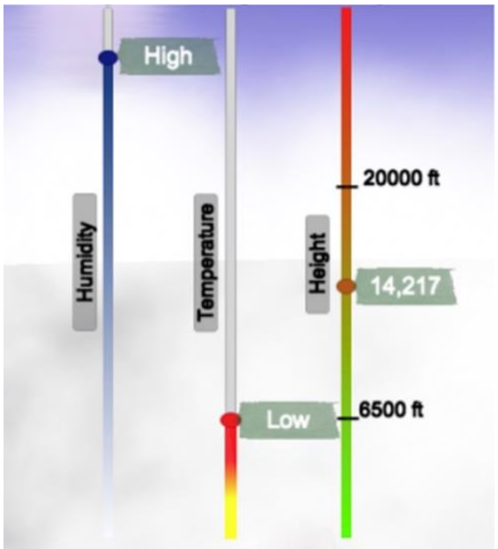
Match the name of the clouds based on the pictures.

	Altostratus	Cumulus	Cirrus
	☒	☐	☐
	☐	☒	☐
	☐	☐	☒

Match the name of the clouds based on their height.

	Low-Level	Mid-Level	High-Level
Cumulus	☒	☐	☐
Altostratus	☐	☒	☐
Cirrus	☐	☐	☒

What is the weather condition if the cloud elevation and temperature and humidity are as bellowed image?



Just Cloud	Rain	Snow
☐	☐	☒

Which type of clouds cover the entire sky?

Cirrus Clouds	Altostratus Clouds	Cumulus Clouds
☐	☒	☐

What fact is true about Cirrus clouds?

They are low and colored gray or white	They look like fluffy	They found very high in the sky	They are the white streaks left behind by airplanes
☐	☐	☒	☐

High level clouds appear 20,000 feet above the surface of Earth.

True	False
☒	☐

Thin, wispy clouds high in the sky are

Altostratus clouds	Cirrus Clouds	Cumulus Clouds
☐	☒	☐

Which of the following is **NOT** a part of weather?

Clouds	Temperature	Tides	Humidity
☐	☐	☒	☐

APPENDIX G: UCF IRB APPROVAL LETTER

UNIVERSITY OF CENTRAL FLORIDA

Institutional Review Board
FWA00000351
IRB00001138, IRB00012110
Office of Research
12201 Research Parkway
Orlando, FL 32826-3246

APPROVAL

April 29, 2020

Dear Eileen Smith:

On 4/29/2020, the IRB reviewed the following submission:

Type of Review:	Initial Study
Title:	Effectiveness of Digital Interactive Experiment in Learning Outcome and Engagement
Investigators:	Eileen Smith and Matin Salemirad
IRB ID:	STUDY00001573
Funding:	None
Grant ID:	None
IND, IDE, or HDE:	None
Documents Reviewed:	• Demographic Survey, Category: Survey / Questionnaire; • Email to School Principal , Category: Recruitment Materials; • Game Engagement Questionnaire , Category: Survey / Questionnaire; • Game Feature Survey , Category: Survey / Questionnaire; • HRP-502b Consent Document (parent for child), Category: Consent Form; • Knowledge Questionnaire (Pre/Post test), Category: Survey / Questionnaire; • Permission Letter to Parents, Category: Recruitment Materials; • rib_HRP-503-Protocol, Category: IRB Protocol;

The IRB approved the protocol on 4/29/2020.

In conducting this protocol, you are required to follow the requirements listed in the Investigator Manual (HRP-103), which can be found by navigating to the IRB Library within the IRB system. Guidance on submitting Modifications and a Continuing Review or Administrative Check-in are detailed in the manual. When you have completed your research, please submit a Study Closure request so that IRB records will be accurate.

Institutional Review Board
FWA00000351
IRB00001138
Office of Research
12201 Research Parkway
Orlando, FL 32826-3246

UNIVERSITY OF CENTRAL FLORIDA

CLOSURE

October 9, 2020

Dear Eileen Smith:

On 10/9/2020, the IRB reviewed the following protocol:

Type of Review:	Continuing Review
Title:	Effectiveness of Digital Interactive Experiment in Learning Outcome and Engagement
Investigators:	Eileen Smith and Matin Salemirad
IRB ID:	CR00000780
Funding:	None
Grant ID:	None
IND, IDE, or HDE:	None

The IRB acknowledges your request for closure of the protocol effective as of 10/9/2020. As part of this action:

- The protocol is permanently closed to enrollment.
- All subjects have completed all protocol-related interventions.
- Collection of private identifiable information is completed.
- Analysis of private identifiable information is completed.

If you have any questions, please contact the UCF IRB at 407-823-2901 or irb@ucf.edu. Please include your project title and IRB number in all correspondence with this office.

Sincerely,

Gillian Bernal
Designated Reviewer

LIST OF REFERENCES

Abt, C. C. (1970). *Serious games*. New York: Viking Press.

Adams, W. K., Paulson, A., & Wieman, C. E. (2008). *What levels of guidance promote engaged exploration with interactive simulations?* Paper presented at the AIP conference proceedings.

Agarwal, R., & Karahanna, E. (2000). Time Flies When You're Having Fun: Cognitive Absorption and Beliefs about Information Technology Usage. *MIS Quarterly, 24*(4), 665–694.

Aghababyan, A. (2014). *E3: Emotions, Engagement, and Educational*. Retrieved from Utah State University, 4031:

Ahmad, M. (2019). Categorizing Game Design Elements into Educational Game Design Fundamentals. In *Game Design and Intelligent Interaction*: IntechOpen.

Ahmad, M., Rahim, L. A., Osman, K., & Arshad, M. I. (2019). Toward Modeling Effective Educational Games Using Mult-Domain Framework. In *Advanced Methodologies and Technologies in Media and Communications* (pp. 165-175). Hershey, PA: IGI Global.

Akaygun, S., & Jones, L. L. (2014). How Does Level of Guidance Affect Understanding When Students Use a Dynamic Simulation of Liquid–Vapor Equilibrium? In *Learning with Understanding in the Chemistry Classroom* (pp. 243-263). Netherlands: Springer.

Aldrich, C. (2009). *The complete guide to simulations and serious games: How the most valuable content will be created in the age beyond Gutenberg to Google*: John Wiley & Sons.

Alofs, T., Theune, M. e., & Swartjes, I. (2011). A Tabletop Interactive Storytelling System:Designing for Social Interaction. *International Journal of Arts and Technology, 8*(3), 199-211.

Alves, A., Lopes, R., Matos, P., Velho, L. & Silva, D. (2010). *Reactoon: Storytelling in a.* Paper presented at the 3rd IEEE International Conference on Digital Game and Intelligent Toy Enhanced Learning.

Angeles, J. P. (2015, Apr. 01). [US9600592B2 - Systems, apparatus, and processes for using smart learning tools with touchscreen devices].

Angelini, M. L. (2016). Integration of the Pedagogical Models “Simulation” and “Flipped Classroom” in Teacher Instruction. *SAGE Open, 6*(1), 2158244016636430. doi:10.1177/2158244016636430

Barke, H.-D., Hazari, A., & Yitbarek, S. (2009). *Misconceptions in chemistry: Addressing perceptions in chemical education*.

Barrutia, O., Ruiz-González, A., Domingo Villarroel, J., & Diez, J. R. (2019). Primary and Secondary Students' Understanding of the Rainfall Phenomenon and Related Water Systems: a Comparative Study of Two Methodological Approaches. *Research in Science Education*.

Beatty, A., & Schweingruber, H. (2016). Seeing students learn science: Integrating assessment and instruction in the classroom [Adobe Digital Edition version]. In.

Belgium, L. (2010). *Out of sight, out of mind: co-player effects on seniors' player experience.* Paper presented at the Fun and Games '10 Proceedings of the 3rd International Conference on Fun and Games, Leuven, Belgium.

Bell, A. (2007). Designing and testing questionnaires for children. *Journal of Research in Nursing, 12*(5), 461-469.

Bénabou, R., & Tirole, J. (2003). Intrinsic and Extrinsic Motivation. *The Review of Economic Studies, 70*(3), 489-520. doi:10.1111/1467-937x.00253

Bjork, S., & Holopainen, J. (2004). *Patterns in game design (game development series)*: Charles River Media, Inc.

Block, F., Hammerman, J., Horn, M., Spiegel, A., Christiansen, J., Phillips, B., . . . Shen, C. (2015). *Fluid Grouping: Quantifying Group Engagement around Interactive Tabletop Exhibits in the Wild.* Paper presented at the CHI '15 Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, Seoul, Republic of Korea.

Blythe, M., & Hassenzahl, M. (2005). The Semantics of Fun: Differentiating Enjoyable Eeperiences. In *Funology From Usability to Enjoyment* (pp. 91-100). Dordrecht,Netherlands: Springer.

Borgers, N., & Hox, J. J. (2000). *Reliability of responses in questionnaire research with children plus coding scheme: a technical report.* Paper presented at the 5th International Conference on Logic and Methodology, Cologne, Germany.

Bourgonjon, J., Valcke, M., Soetaert, R., & Schellens, T. (2010). Students' perceptions about the use of video games in the classroom. *Computers & Education, 54*(4), 1145-1156. doi:10.1016/j.compedu.2009.10.022

Bouvier, P., Lavoué, E., & Sehaba, K. (2014). Defining Engagement and Characterizing Engaged-Behaviors in Digital Gaming. *Simulation & Gaming, 45*(4-5), 491-507. doi:10.1177/1046878114553571

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

Brockmyer, J. H., Fox, C. M., Curtiss, K. A., McBroom, E., Burkhart, K. M., & Pidruzny, J. N. (2009). The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology, 45*, 624-634.

Carolyn M. Plump, J. L. (2017). Using Kahoot! in the Classroom to Create Engagement and Active Learning: A Game-Based Technology Solution for eLearning Novices. *Management Teaching Review, 2*(2), 151-158.

Cavus, N., & Ibrahim, D. (2017). Learning English using children's stories in mobile devices. *British Journal of Educational Technology, 48*(2), 625-641.

Chang, C.-C., Liang, C., Chou, P.-N., & Lin, G.-Y. (2017). Is game-based learning better in flow experience and various types of cognitive load than non-game-based learning? Perspective from multimedia and media richness. *Computers in Human Behavior, 71*, 218-227.

Chang, C.-C., Liang, C., Chou, P.-N., & Lina, G.-Y. (2017). Is game-based learning better in flow experience and various types of cognitive load than non-game-based learning? Perspective from multimedia and media richness. *Computers in Human Behavior, 71*, 218-227.

Chang, M., Evans, M., Kim, S., Deater-Deckard, K., & Norton, A. (2014). *Educational video games and Students' game engagement.* Paper presented at the 2014 International Conference on Information Science & Applications (ICISA).

Cheng, M. T., She, H. C., & Annetta, L. A. (2015). Game immersion experience: its hierarchical structure and impact on game-based science learning. *Journal of Computer Assisted Learning, 31*(3), 232-253.

Christenson, S. L., Reschly, A. L., & Wylie, C. (2012). *Handbook of Research on Student Engagement*. London: Springer.

Chu, H.-C. (2014). Potential negative effects of mobile learning on students' learning achievement and cognitive load - a format assessment perspective. *Educational Technology & Society, 17*(1), 332-344.

Clark, A. (1997). *Being There: Putting Brain, Body, and World Together Again*. London, England: The MIT Press.

Compton, K., & Mateas, M. (2015). *Casual Creators.* Paper presented at the ICCC.

Connolly, T. M., A. Boyle, E., MacArthur, E., Hainey, T., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education, 59*(2), 661-686.

Coomans, M. K., & Timmermans, H. J. (1997). *Towards a Taxonomy of Virtual Reality User Interfaces.* Paper presented at the Information Visualization, London.

Csíkszentmihályi, M. (1997). *Flow and the psychology of discovery and invention*. New York: Harper Perennial.

De Kort, Y. A., IJsselsteijn, W. A., & Poels, K. (2007). Digital games as social presence technology: Development of the Social Presence in Gaming Questionnaire (SPGQ). *Proceedings of PRESENCE, 195203*.

De Lope, R. P., Medina-Medina, N., Soldado, R. M., García, A. M., & Gutiérrez-Vela, F. L. (2017). *Designing educational games: Key elements and methodological approach.* Paper presented at the 2017 9th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games).

Delaney, M. L., & Royal, M. A. (2017). Breaking Engagement Apart: The Role of Intrinsic and Extrinsic Motivation in Engagement Strategies. *Industrial and Organizational Psychology, 10*(1), 127-140.

Delialioğlu, Ö. (2012). Student Engagement in Blended Learning Environments with Lecture-Based and Problem-Based Instructional Approaches. *Educational Technology & Society, 15*(3), 310-322.

Denny, P. (2013). *The Effect of Virtual Achievements on Student Engagement.* Paper presented at the CHI '13 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Paris, France.

Derick Kiger, D. H., Deb Prunty. (2012). Examining the Influence of a Mobile Learning Intervention on Third Grade Math Achievement. *Journal of Research on Technology in Education, 45*(1), 61-82.

Domínguez, A., Saenz-de-Navarrete, J., de-Marcos, L., Fernández-Sanz, L., Pagés, C., & Martínez-Herráiz, J.-J. (2013). Gamifying learning experiences: Practical implications and outcomes. *Computers & Education, 63*, 380-392.

Earth CARE - Earth Online. Retrieved from https://earth.esa.int/web/guest/missions/esa-future-missions/earthcare

Echeverría, A., Barrios, E., Nussbaum, M., Améstica, M., & Leclerc, S. (2012). The atomic intrinsic integration approach: A structured methodology for the design of games for the conceptual understanding of physics. *Computers & Education, 59*(2), 806-816.

Ecker, M., Müller, W., & Zylka, J. (2011). Game-based learning design patterns: An approach to support the development of “better” educational games. In *Handbook of research on improving learning and motivation through educational games: Multidisciplinary approaches* (pp. 137-152): IGI Global.

Eleni A. Kyza, S. E., & Tiberghien, A. (2009). Technology-Enhanced Learning in Science. In *Technology-Enhanced Learning* (pp. 121-134). 2009: Springer,Dordrecht.

Emmanuel Tsekleves, J. C., Amar Aggoun. (2016). Benefits, barriers and guideline recommendations for the implementation of serious games in education for stakeholders and policymakers. *British Journal of Educational Technology, 47*(1), 164-183.

Filseckera, M., & Hickey, D. T. (2014). A multilevel analysis of the effects of external rewards on elementary students' motivation, engagement and learning in an educational game. *Computers & Education, 75*, 136-148.

Flores, N., Paiva, A. C., & Cruz, N. (2020). Teaching Software Engineering Topics Through Pedagogical Game Design Patterns: An Empirical Study. *Information, 11*(3), 153.

Foster, J. A., Sheridan, P. K., Irish, R., & Frost, G. S. (2012). *Gamification as a strategy for promoting deeper investigation in a reverse engineering activity.* Paper presented at the Proceedings of the 2012 American Society for Engineering Education Conference.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School Engagement: Potential of the Concept, State of the Evidentce. *Review of Educational Research, 74*(1), 59-109.

Freitas, S. d. (2018). Are Games Effective Learning Tools? A Review of Educational Games. *Journal of Educational Technology & Society, 21*(2), 74–84.

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & Gaming, 33*(4), 441-467.

Gee, J. P. (2008). Learning and Games. In *The Ecology of Games: Connecting Youth, Games, and Learning* (pp. 21-40). Cambridge, MA.: The MIT Press.

Gee, J. P. (2009). Deep Learning Properties of Good Digital Games: How Far Can They Go? In *Serious Games: Mechanisms and Effects* (pp. 89*104). New York: Routledge.

Gorrall, B. K., Curtis, J. D., Little, T. D., & Panko, P. (2016). Innovations in Measurement: Visual Analog Scales and Retrospective Pretest Self-Report Designs. *Actualidades en Psicologí, 30*(120), 1-6.

Guin, T. D.-L., Baker, R., Mechling, J., & Ruyle, E. (2012). Myths and realities of respondent engagement in online surveys. *International Journal of Market Research, 54*(5), 613-633.

Gurria, A. (2015). *PISA 2015 Results in Focus*. Retrieved from Paris:

Hainey, T., Connolly, T. M., Boyle, E. A., Wilson, A., & Razak, A. (2016). A systematic literature review of games-based learning empirical evidence in primary education. *Computers & Education, 102*(C), 202-223.

Hall, L., Hume, C., & Tazzyman, S. (2016). *Five Degrees of Happiness: Effective Smiley Face Likert Scales for Evaluating with Children.* Paper presented at the The 15th International Conference.

Hamari, J., & Koivisto, J. (2014). Measuring flow in gamification: Dispositional Flow Scale-2. *Computers in Human Behavior, 40*, 133-143. doi:https://doi.org/10.1016/j.chb.2014.07.048

Hamari, J., Shernoff, D. J., Rowe, E., Edwards, T., Asbell-Clarke, J., & Asbell-Clarke, J. (2016). Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning. *Computers in Human Behavior, 54*, 170-179.

Hawlitschek, A., & Joeckel, S. (2017). Increasing the effectiveness of digital educational games: The effects of a learning instruction on students' learning, motivation and cognitive load. *Computers in Human Behavior, 72*, 79-86. doi:10.1016/j.chb.2017.01.040

Hayes, M. H., & Patterson, D. G. (1921). Experimental development of the graphic rating method. *Psychological Bulletin, 18*, 98-99.

Hess, T., & Gunter, G. (2013). Serious game-based and nongame-based online courses:Learning experiences and outcomes. *British Journal of Educational Technology, 44*(3), 372-385.

Hoffman, B., & Nadelson, L. (2010). Motivational engagement and video gaming: A mixed methods study. *Educational Technology Research and Development, 58*(3), 245-270.

Honey, M. A., & Hilton, M. (2011). *Learning Science Through Computer Games and Simulations*. Retrieved from Washington, D.C.:

Huizenga, J., Admiraal, W., Akkerman, S., & Dam, G. t. (2009). Mobile game-based learning in secondary education: engagement, motivation and learning in a mobile city game. *Journal of Computer Assisted Learning, 25*(4), 332-344.

Huynh-Kim-Bang, B., Wisdom, J., & Labat, J.-M. (2010). Design Patterns in Serious Games: A Blue Print for Combining Fun and Learning. *Project SE-SG*.

IJsselsteijn, W. A., de Kort, Y.A.W. & Poels, K. . (2013). The Game Experience Questionnaire. Eindhoven:

Technische Universiteit Eindhoven. *FUGA The fun of gaming: Measuring the human experience of media enjoyment*.

Iten, N., & Petko, D. (2016). Learning with serious games: Is fun playing the game apredictor of learning success? *British Journal of Educational Technology, 47*(1), 151-163.

Izar, C., Stark, K., Trentacosta, C. J., & Schultz, D. (2008). Beyond Emotion Regulation: Emotion Utilization and Adaptive Functioning. *Child Development Perspectives*, 156-163.

Jabbar, A. I. A., & Felicia, P. (2015). Gameplay Engagement and Learning in Game-Based Learning. *Waterford Institute of Technology, 85*(4), 740-779.

Jackson, S. A., Martin, A. J., & Eklund, R. C. (2008). Long and short measures of flow: The construct validity of the FSS-2, DFS-2, and new brief counterparts. *Journal of Sport and Exercise Psychology, 30*(5), 561-587.

Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies, 66*(9), 641-661.

Jenny Eppard, O. N., Preeya Reddy. (2016). The Next Generation of Technology: Mobile Apps in the English Language Classroom. *International Journal of Emerging Technologies in Learning, 11*(4), 21-27.

Jimoyiannis, A. (2009). Computer Simulations and Scientific Knowledge Construction. In *Encyclopedia of Information Communication Technology* (pp. 15): IGI-Global.

Johnson, C. (Producer). (2018, Aug. 13). Understanding the 6 Types of Response Bias (With Examples). *www.nextiva.com*. Retrieved from https://www.nextiva.com/blog/response-bias.html

K. D'Mello, S., & C. Graesser, A. (2014). Confusion. *International handbook of emotions in education*, 299-320.

Ke, F., & Abras, T. (2013). Games for engaged learning of middle school children with special learning needs. *British Journal of Educational Technology, 44*(2), 225-242. doi:10.1111/j.1467-8535.2012.01326.x

Kebritchi, M. (2008). *Effects Of A Computer Game On Mathematics Achievement And Class Motivation: An Experimental Study / Unpublished doctoral dissertation*. Retrieved from Orlando, Florida:

Kenny, R. F., & Mcdaniel, R. (2011). The role teachers' expectations and value assessments of video games play in their adopting and integrating them into their classrooms. *British Journal of Educational Technology, 42*(2), 197 - 213.

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

Kiilia, K., Freitas, S. d., Arnab, S., & Lainema, T. (2012). The Design Principles for Flow Experience in Educational. *Procedia Computer Science, 15*, 78-91.

Kim, S., Chang, M., Deater-Deckard, K., Evans, M. A., Norton, A., & Samur, Y. (2017). Educational games and students' game engagement in elementary school classrooms. *Computers in Education, 4*(4), 395–418.

Koster, R. (2013). *A Theory of Fun for Game Design*. Scottsdale, Arizona: O'Reilly Media, Inc.

Lameras, P., Arnab, S., Dunwell, I., Stewart, C., Clarke, S., & Petridis, P. (2017a). Essential features of serious games design in higher education: Linking learning attributes to game mechanics. *British Journal of Educational Technology, 48*(4), 972-994. doi:10.1111/bjet.12467

Lameras, P., Arnab, S., Dunwell, I., Stewart, C., Clarke, S., & Petridis, P. J. B. j. o. e. t. (2017b). Essential features of serious games design in higher education: Linking learning attributes to game mechanics. *48*(4), 972-994.

Lazo, J. K., Morss, R. E., & Demuth, J. L. (2009). 300 Billion Served: Sources, Perceptions, Uses, and Values of Weather Forecasts. *Bulletin of the American Meteorological Society, 90*(6), 785-798.

Li, Q., Lemieux, C., Vandermeiden, E., & Nathoo, S. (2013). Are you ready to teach secondary mathematics in the 21st century? A study of preservice teachers' digital game design experience. *Journal of Research on Technology in Education, 45*(4), 309-337.

Lieberman, D. A., Fisk, M. C., & Biely, E. (2009). Digital Games for Young Children Ages Three to Six: From Research to Design. *Computers in the Schools, 26*(4), 299-313. doi:10.1080/07380560903360178

Linke, R., Kothe, T., & Alt, F. (2017). *TaBooGa: A Hybrid Learning App to Support Children's Reading Motivation.* Paper presented at the Conference on Interaction Design and Children, Stanford, California, USA.

Liou, W.-K., Bhagat, K. K., & Chang, C.-Y. (2018). The design, implementation, and evaluation of a digital interactive globe system integrated into an Earth Science course. *Educational Technology Research and Development, 66*(2), 545-561.

Lombard, M., Weinstein, L., & Ditton, T. (2011). *Measuring telepresence: The validity of the Temple Presence Inventory (TPI) in a gaming context.* Paper presented at the ISPR 2011: The International Society for Presence Research Annual Conference.

Luebeck, J., Roscoe, M., Cobbs, G., Diemert, K., & Scott, L. (2017). Re-envisioning Professional Learning in Mathematics: Teachers' Performance, Perceptions, and Practices in Blended Professional Development. *Journal of Technology and Teacher Education, 25*(3), 273-299.

M.Connolly, T., A.Boyle, E., MacArthur, E., Hainey, T., & M.Boyle, J. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education, 59*(2), 661-686.

Maleus, E., Kilas, E., & Kruus, S. (2016). Students' understanding of clouds and rainbow formation and teachers' awareness od students' performance. *International Journal of Science Education, 38*(6), 993-1011.

Marks, H. M. (2000). Student Engagement in Instructional Activity: Patterns in the Elementary, Middle, and High School Years. *American Educational Research, 37*(1), 153-184.

Mayes, D. K., & Cotton, J. E. (2001). Measuring Engagement in Video Games: A Questionnaire. *Proceedings of the Human Factors & Ergonomics Society Annual Meeting*, 692-696. Retrieved from https://login.ezproxy.net.ucf.edu/login?auth=shibb&url=https://search.ebscohost.com/login.aspx?direct=true&db=ega&AN=ega182953&site=eds-live&scope=site

McGonigal, J. (2011). *Reality is broken: Why games make us better and how they can change the world*: Penguin.

Merhi, O., Faugloire, E., Flanagan, M., & Stoffregen, T. A. (2007). Motion Sickness, Console Video Games, and Head-Mounted Displays. *Human Factors, 49*(5), 920-934.

Min Liu, C. C. N., Jennifer Wivagg. (2014). Potentials of Mobile Technology for K-12 Education: An Investigation of iPod touch Use for English Language Learners in the United States. *Journal of Educational Technology & Society, 17*(2), 115-126.

Moreno, R., & Mayer, R. E. (2005). Role of guidance, reflection, and interactivity in an agent-based multimedia game. *Journal of Educational Psychology, 97*(1), 117.

NASA - ACE. Retrieved from https://acemission.gsfc.nasa.gov/

. National Academies of Sciences, Engineering, and Medicine. (2019). *Science and Engineering for Grades 6-12: Investigation and*
Design at the Center.
New Schools 2018-19. (2018). *OCPS (Orange County Public Schools)*. Retrieved from https://www.ocps.net/departments/facilities/facilities_highlights
Newmann, F. M., Wehlage, G. G., & Lamborn, S. D. (1992). The Significance and Sources of Student Engagement. In *Student Engagement and Achievement in American Secondary Schools* (pp. 11-39). New York: Teachers College, Columbia University.
NGSSS. (2019). Weather and Climate. *Next Generation Science Standards*. Retrieved from https://www.nextgenscience.org/topic-arrangement/kweather-and-climate
Nicholls, M. E. R., Orr, C. A., Okubo, M., & Loftus, A. (2006). Satisfaction Guaranteed: The Effect of Spatial Biases on Responses to Likert Scales. *Psychological Science, 17*(13), 1027-1028.
Nicoll, P., MacRury, S., Woerden, H. C. v., & Smyth, K. (2018). Evaluation of Technology-Enhanced Learning Programs for Health Care Professionals: Systematic Review. *Journal of Medical Internet Research, 20*(4), e131.
O'Brien, H. L., & Toms, E. G. (2008). What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology, 59*(6), 938-955.
O'Brien, H. L., & Toms, E. G. (2010). The development and evaluation of a survey to measure user engagement. *Journal of the American Society for Information Science and Technology, 61*(1), 50-69.
O'Muircheartaigh, C., Gaskell, G., & Wright, D. B. (1995). Weighing Anchors: Verbal and Numeric Labels for Response Scales. *Journal of Official Statistics, 11*(3), 295-307.
OECD. (2016). *Education at a Glance 2016, OECD Indicators*. Retrieved from Paris:
Perski, O., Blandford, A., West, R., & Michie, S. (2017). Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. *Transl Behav Med, 7*(2), 254-267. doi:10.1007/s13142-016-0453-1
Pesare, E., Roselli, T., Corriero, N., & Rossano, V. J. S. L. E. (2016). Game-based learning and Gamification to promote engagement and motivation in medical learning contexts. *3*(1), 5. doi:10.1186/s40561-016-0028-0
Phan, M. d., Montz, B. e., Curtis, S., & RiCkenbaCh, T. M. (2018). Weather on the Go: An Assessment of Smartphone Mobile Weather Application Use among College Students. *Bulletin of the American Meteorological Society, 99*(11), 2245-2257.
PhET (Producer). (2019, May 05). PhET: Research and Development. *PhET Interactive Simulations*. Retrieved from https://phet.colorado.edu/en/research
Prensky, M. (2003). Digital game-based learning. *Computers in Entertainment (CIE), 1*(1), 1-4.
Prensky, M. (2003). Digital Game-Based Learning. *ACM Computers in Entertainment, 1*(1), 21-21.
Psaltis, A., Apostolakis, K. C., Dimitropoulos, K., & Daras, P. (2018). Multimodal student engagement recognition in prosocial games. *IEEE Transactions on Games, 10*(3), 292-303.
Pundak, D., Liberman, I., & Shacham, M. (2017). From Conceptual Frameworks to Mental Models for Astronomy: Students' Perceptions. *Journal of Astronomy & Earth Sciences Education, 4*(2), 109-126.
Quinn, C. N. (2005). *Engaging Learning: Designing e-Learning Simulation Games*. San Francisco, CA.: John Wiley & Sons.

Räsänen, P., Laurillard, D., Käse, T., & Aster, r. v. (2019). Perspectives to Technology-Enhanced Learning and Teaching in Mathematical Learning Difficulties. *International Handbook of Mathematical Learning Difficulties*, 733-754.

Ravaja, N., Saari, T., Salminen, M., Laarni, J., & Kallinen, K. (2006). Phasic Emotional Reactions to Video Game Events: A Psychophysiological Investigation. *Media Psychology, 8*(4), 343-367. doi:10.1207/s1532785xmep0804_2

Reeve, J., & Tseng, C.-M. (2011). Agency as a fourth aspect of students' engagement during learning activities. *Contemporary Educational Psychology, 36*(4), 257-267.

Reynolds-Keefer, L., & Johnson, R. (2011). Is a Picture Is Worth a Thousand Words? Creating Effective Questionnaires with Pictures. *Practical Assessment, Research & Evaluation, 16*(8), 1-7.

Rideout, V. (2016). Measuring time spent with media: the Common Sense census of media use by US 8- to 18-year-olds. *Journal of Children and Media, 10*(1), 138-144.

Riedl, M. O., & Bulitko, V. (2013). Interactive narrative: An intelligent systems approach. *Ai Magazine, 34*(1), 67-67.

Riedl, M. O., & Bulitko, V. J. A. M. (2013). Interactive narrative: An intelligent systems approach. *34*(1), 67-67.

Roberts, J., Banerjee, A., Hong, A., McGee, S., Horn, M., & Matcuk, M. (2018). *Digital Exhibit Labels in Museums: Promoting Visitor Engagement with Cultural Artifacts.* Paper presented at the 2018 HCI Conference on Human Factors in Computing Systems, Montréal, QC, Canada.

Ronimusac, M., Kujalab, J., Tolvanenc, A., & Lyytinen, H. (2014). Children's engagement during digital game-based learning of reading: The effects of time, rewards, and challenge. *Computers & Education, 71*, 237-246.

Rooney, P. (2012). A theoretical framework for serious game design: exploring pedagogy, play and fidelity and their implications for the design process. *International Journal of Game-Based Learning (IJGBL), 2*(4), 41-60.

Rosas, R., Nussbaum, M., Cumsille, P., Marianov, V., Correa, M. n., Flores, P., . . . Salinas, M. (2003). Beyond Nintendo: design and assessment of educational video games for first and second grade students. *Computers & Education, 40*, 71-94.

Rozendaal, M. C., Braat, B. A. L., & Wensveen, S. A. G. (2010). Exploring sociality and engagement in play through game-control distribution. *AI & Society, 25*(2), 193-201.

Rutten, N., Van Joolingen, W. R., & Van Der Veen, J. J. (2012). The learning effects of computer simulations in science education. *58*(1), 136-153.

S. Gould, K., Gilbert, A., J. Pike, A., & J. Menzies, I. (2018). Interactive touch-screen monitors facilitate collaborative learning of microscopy skills in an introductory-level plant biology lab. *Journal of Biological Education*, 1-7.

Sadler, T. D., Romine, W. L., Stuart, P. E., & Merle-Johnson, D. (2013). Game-Based Curricula in Biology Classes: Differential Effects Among Varying Academic Levels. *Journal of Research in Science Teaching, 50*(4), 479-499.

Salen, K., & Zimmerman, E. (2004). *Rules of Play : Game Design Fundamentals*: The MIT Press.

Sauro, J. (Producer). (2010, Sep. 14). Survey Respondents Prefer the Left Side of a Rating Scale. *measuringu.com*. Retrieved from https://measuringu.com/left-side-bias/

Schell, J. (2008). *The art of game design:a book of lenses*. San Francisco, CA: Morgan Kaufmann Publishers Inc.

Schwartz, D. L., Bransford, J. D., & Sears, D. J. T. o. l. f. a. m. m. p. (2005). Efficiency and innovation in transfer. *3*, 1-51.

Sengupta, P., & Clark, D. (2016). Playing modeling games in the science classroom: The case for disciplinary integration. *Educational Technology, 56*(3), 16 - 22.
Shi, Y.-R., & Shih, J.-L. (2015). Game Factors and Game-Based Learning Design Model. *International Journal of Computer Games Technology*, 11.
Shute, V. J., & Ke, F. (2012). Games, Learning, and Assessment. *Assessment in Game-Based Learning*, 43-58.
Simba Information. (2016). Retrieved from Stanford, CT.:
Sinatra, G. M., Heddy, B. C., & Doug, L. (2015). The Challenges of Defining and Measuring Student Engagement in Science. *Educational Psychologist, 50*(1), 1-13.
Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The Challenges of Defining and Measuring Student Engagement in Science. *Educational Psychologist, 501*, 1-13.
Sitzmann, T. J. P. p. (2011). A meta-analytic examination of the instructional effectiveness of computer-based simulation games. *64*(2), 489-528.
Skew-T Log-P Diagrams. (2019). Retrieved from https://www.weather.gov/jetstream/skewt
Skinner, E., Furrer, C., Marchand, G., & Kindermann, T. (2008). Engagement and Disaffection in the Classroom: Part of a Larger Motivational Dynamic? *Educational Psychology, 100*(1), 765-781.
Skinner, E. A., & Belmont, M. J. (1993). Motivation in the Classroom: Reciprocal Effects of Teacher Behavior and Student Engagement Across the School Yea. *Journal of Educational Psychology, 85*(4), 571-581.
Soper, H., Young, A., Cave, B., Lee, A., & Pearson, K. (1917). On the distribution of the correlation coefficient in small samples. Appendix II to the papers of" Student" and RA Fisher. *Biometrika, 11*(4), 328-413.
Stange, M., Barry, A., Smyth, J., & Olson, K. (2016). Effects of Smiley Face Scales on Visual Processing of Satisfaction Questions in Web Surveys. *Social Science Computer Review*, 1-11.
Swartout, W., & van Lent, M. (2003). Making a game of system design. *Communications of the ACM, 46*(7), 32-39.
Tan, J. L., Goh, D. H.-L., P.Ang, R., & S.Huan, V. (2013). Participatory evaluation of an educational game for social skills acquisition. *Computers & Education, 64*, 70-80.
Ting, Y.-L. (2012). The pitfalls of mobile devices in learning: A different view and implications for pedagogical design. *Journal of Educational Computing Research, 46*(2), 119-134.
Tsekleves, E., Cosmas, J., & Aggoun, A. J. B. J. o. E. T. (2016). Benefits, barriers and guideline recommendations for the implementation of serious games in education for stakeholders and policymakers. *47*(1), 164-183.
UCAR. Retrieved from http://tornado.sfsu.edu/geosciences/classes/m201/buoyancy/SkewTMastery/mesoprim/skewt/description.htm
UCP 2019: Understanding Clouds and Precipitation | GEWEX Events. (2019). Retrieved from https://www.gewexevents.org/events/ucp-2019-understanding-clouds-and-precipitation/
Villarroel, J. D., & Ros, I. (2013). Young Children's Conceptions of Rainfall:A study of Their Oral and Pictoral Explanations. *International Education Studies, 6*(8).
Viloria, A., Lezama, O. B. P., & Cabrera, D. (2020). Virtual reality tool to improve collaborative work in the university context. *Procedia Computer Science, 175*, 635-640.
Vos, N., Meijden, H. v. d., & Denessen, E. (2011). Effects of constructing versus playing an educational game on student motivation and deep learning strategy use. *Computers & Education, 50*(1), 127-137.

Vosniadou, S., & Skopeliti, I. (2017). Is it the Earth that turns or the Sun that goes behind the mountains? Students' misconceptions about the day/night cycle after reading a science text. *International Journal of Science Education, 39*(15), 2027-2051.
Wang, M.-T., Willett, J. B., & Eccles, J. S. (2011). The assessment of school engagement: Examining dimensionality and measurement invariance by gender and race/ethnicity. *Journal of School Psychology, 49*(4), 465-480.
Wang, M. T., & Eccles, J. S. (2012). Adolescent Behavioral, Emotional, and Cognitive Engagement Trajectories in School and Their Differential Relations to Educational Success. *Research on Adolescence, 22*(1), 31-39.
Weber, R., Tamborini, R., Westcott-Baker, A., & Kantor, B. (2009). Theorizing Flow and Media Enjoyment as Cognitive Synchronization of Attentional and Reward Networks. *Communication Theory, 19*, 397–422.
Weibel, D., & Wissmath, B. (2011). Immersion in Computer Games: The Role of Spatial Presence and Flow. *International Journal of Computer Games Technology*, 1-14.
Wentzel, K. R., & Miele, D. B. (2009). *Handbook of motivation at school*: Routledge.
Whitton, N. (2007). *Motivation and computer game based learning.* Paper presented at the Proceedings of the Australian Society for Computers in Learning in Tertiary Education, Singapore.
Wiebe, E. N., Lamb, A., Hardy, M., & Sharek, D. (2014). Measuring engagement in video game-based environments: Investigation of the User Engagement Scale. *Computers in Human Behavior, 32*, 123-132.
Wilcoxon, F. (1992). Individual comparisons by ranking methods. In *Breakthroughs in statistics* (pp. 196-202): Springer.
Willms, J. D. (2003). *STUDENT ENGAGEMENT AT SCHOOL A SENSE OF BELONGING AND PARTICIPATION*. Paris: Organisation for Economic Co-Operation and Development.
Witmer, B. G., & Singer, M. J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence, 7*(3), 225-240.
Yang, W., Rifqi, M., Marsala, C., & Pinna, A. (2018). *Towards Better Understanding of Player's Game Experience*. Paper presented at the Proceedings of the 2018 ACM on International Conference on Multimedia Retrieval - ICMR '18.
Zhang, T., Chen, A., & Ennis, C. (2017). Elementary school students' naïve conceptions and misconceptions about energy in physical education context. *Sport, Education and Society, 24*(1), 25-37.
Zhu, Y., Pei, L., & Shang, J. (2017). *Improving Video Engagement by Gamification: A Proposed Design of MOOC Videos.* Paper presented at the International Conference on Blending Learning (ICBL), Cham.
Zulehmay, M. N., Tarmizi, M. L. M., Rahid, N. N. A., Afiqah, A., Sharip, N. F. M. Z., & Johari, N. F. (2019). Optimizing Gamification in Learning Solar System for Primary School.