Lehigh University
Lehigh Preserve

Theses and Dissertations

2013

The Effects of Digital Games on Middle School Students' Mathematical Achievement

Pilar Lisa Starkey
Lehigh University

Follow this and additional works at: http://preserve.lehigh.edu/etd

Part of the Educational Leadership Commons

Recommended Citation
Starkey, Pilar Lisa, "The Effects of Digital Games on Middle School Students' Mathematical Achievement" (2013). *Theses and Dissertations.* Paper 1637.

This Dissertation is brought to you for free and open access by Lehigh Preserve. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Lehigh Preserve. For more information, please contact preserve@lehigh.edu.

The Effects of Digital Games on Middle School Students' Mathematical Achievement

by

Pilar Lisa Starkey

Presented to Dissertation Committee

of Lehigh University

Dissertation Proposal in Candidacy for the Degree of

Doctor of Education

in

Educational Leadership

Advisor: Dr. Floyd Beachum

Committee Members:

Dr. Thomas Hammond

Dr. Kelly Mekdeci

Dr. Arnold Spokane

Copyright by Pilar Lisa Starkey

2013

This thesis proposal is accepted and approved in partial fulfillment of the requirements for the degree of Doctor of Education.

(Date)

Floyd Beachum, Advisor

Thomas Hammond

Kelly Mekdeci

Arnold Spokane

ACKNOWLEDGEMENTS

I would never have completed this dissertation without the love and support of my family. Bill, your unwavering patience and belief in me means more than words can express. Nahita, Zach, and Catie, you remind me every day of what matters in this world, including the importance of the teaching profession.

Thank you to my dissertation chair, Floyd Beachum for his oversight and insight. Arnold Spokane and Thomas Hammond, thank you for your care, critical feedback, and thoughtfulness. Kelly Mekdeci, we began this journey as unlikely suitemates and have travelled through life's upheavals and successes together. Thank you for being a colleague, mentor, and friend. To Daphne Hobson, without your wisdom and occasional New York hollering, I would never have completed Lehigh's program. Laura Roberts, thank you for assistance with the statistical analysis. So many other colleagues and friends helped along the way – I'll always be grateful for your contributions in helping me achieve this once-in-a-lifetime goal.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract

In the last few years digital games have gained attention as potential tools for facilitating learning in different sectors of society including but not limited to military, health, and education. However, relatively few empirical studies have investigated the effects of digital games in the context of formal K-12 settings. This study examined data collected during a program evaluation to explore the effects of a digital game on middle school male and female students' mathematics achievement, situational motivation, and attitudes toward mathematics. The study included data from 168 students attending a private international school in Africa, who were assigned to treatment and control groups by stratified random sampling to ensure a balance of boys and girls as well as equal representation of students from grade six, seven and eight. Achievement was measured using internal school exams based on benchmarks aligned with the National Council of Teachers of Mathematics (NCTM) standards and benchmarks. Motivation to learn mathematics was measured using the Course Interest Survey (CIS) based on Keller's ARCS model of motivation. Mathematical attitude was measured using the Fennema-Sherman Mathematical Attitude Scales (FSMAS). A Multivariate Analysis of Covariance (MANCOVA) was performed to analyze change from pre- to post-test scores in achievement, motivation, and attitudes with the independent variables of group (control and treatment) and sex (male and female). Results showed a significant increase in mathematical achievement (b = -1.87, p < .0005, ES = .13), motivation to learn in math class (b = -1.17, p < .0005, ES = .42), and attitudes toward mathematics (b = -.77, p < .05, ES = .09; b = -1.18, p < .0005, ES = .13) for both boys and girls who played the digital game.

Keywords: digital games, technology, education, middle school, mathematics

CHAPTER I

Introduction and Rationale

"Are we going to do anything fun today?" Students have posed this question to their teachers for generations, yet educators rarely address it. Typically, secondary school teachers remain focused on prescribed content areas and pay little attention to supporting students in developing an emotional investment in the learning objectives (Appleton, Christenson, & Furlong, 2008). Despite changes in virtually every aspect of life and society over the past hundred years, journalists purport, "[students] spend much of the day as their great-grandparents once did: sitting in rows, listening to teachers lecture, scribbling notes by hand, reading from textbooks that are out of date by the time they are printed. A yawning chasm (with an emphasis on yawning) separates the world inside the schoolhouse from the world outside" (Wallis, Steptoe, & Miranda, 2006, p. 50).

Every generation is different from the previous and some generational gaps have been relatively small. However, exponential technology changes are resulting in widening generational gaps. For example, communication today is vastly different from just a few years ago and the language of our culture is changing to reflect this evolving revolution. Prensky (2001) pointed out that today's students are growing up in a world of social networks, wireless communication, and instant text messages, thus are native speakers of the language of modern technology. The Pew Internet & American Life project's 2002 investigation of students' use of the Internet, *The Digital Disconnect: The Widening Gap Between Internet Savvy Students and Their Schools*, found that "many schools and teachers have not yet recognized—much less responded to—the new ways students communicate and access information" (Levin & Arafeh, 2003, p. iii). While some educators Google, twitter, will BRB, and sometimes ;-) most often they lack the technology fluency of today's students who are "digital natives" born into the current high tech environment (Prensky, 2001). The generational gap between students and the educational system does not just exist in terms of communication, but also in the manner and location in which students are learning to solve problems, explore, and question the world around them (Simpson, 2005). Recent technology developments have transformed society and become integral to people's reading, writing, calculating, and thinking; skills that continue to form the foundation of most schools' curricula. Yet, many recent technologies remain on the periphery of schools (Collins & Halverson, 2009; McLeod, 2011; Simpson, 2005). Prensky contended that to bridge this disconnect, educators, most of whom are "digital immigrants," must adapt to the digital learners' world to effectively reach modern students (2001).

From large American cities to villages in developing nations, emerging technologies are altering the trajectories of change in our world. In 2005, the *New York Times* proclaimed: "Cellphones Catapult Rural Africa to 21st Century" (LaFraniere, 2005). By 2012, more than 50% of citizens in African nations were connected via cell phones and using them for mobile banking, social activism, and games (Ogunlesi & Busari, 2012; Perry, 2011). Even in villages without electricity, running water, or access to modern health care, people are accessing social networking sites. If Facebook were a nation, its more than a billion users (Facebook, 2013) would make it the third most populous country in the world. In the United States, 47% of adults reported getting their news via their cell phones or computers (Purcell, Rainie, Rosenstiel, & Mitchell, 2011). Therefore, it is not surprising that the Annenberg School for Communication & Journalism at the University of Southern California predicts, "most printed daily newspapers will be gone in about five years" (The Center for the Digital Future, 2012, p. 13). As of 2012, Internet World Stats reported that over 2.4 billion people used the Internet. A 2008 study conducted for the Pew Internet and American Life Project found 97% of teens aged 12-17 played some form of electronic game, (Lenhart et al., 2008). Collectively, digital gamers spent over 40 billion dollars in 2010 (Chatfield, 2010). McGonigal estimated that collectively *World of Warcraft* players have spent more years, nearly six million, in the game than humans have been on Earth (2011).

The ways in which people can communicate, share, create, make friends, and learn is changing at a remarkable rate. Technology has begun to have an effect on the classroom structure; however most institutions are still unable to make full use of the benefits it can bring. The vast majority of schools still adhere to a 19th century factory (industrial) model of education in which all students are taught the same material from the same textbook in the same manner at the same time (Schrenko, 1994; Darling-Hammond et al., 2009). “It [the educational system] is frozen in time, based on assumptions that don't fit the current world. We need a broader vision of what it means to educate, not just how to integrate technology. This means we need to redefine what it means to both teach and learn” (Anderson, Boyles, & Rainie, 2012, p. 28). Despite the richness of options technology innovations provide for communicating, learning, and creating, most often students have to “power down” when they come to class (Prensky, 2006, p. 10).

Digital games offer “a new pedagogy for the 21st century, one that has the potential to not merely fill individual minds, but empower whole persons, and to transform learning from a rote acquisitional process to a transactive one in which conceptual understandings have transformational significance” (Barab, Gresalfi, Dodge, & Ingram-Goble, 2010). Most young people in the education system today have grown up in a world where digital games are a part of life, yet these tools are not actively embraced by school systems that are still entrenched in an antiquated model of education (Chatfield, 2010; Klopfer, Osterweil, & Salen, 2009).

Despite the rhetoric calling for pedagogical changes based on research about how students learn and the opportunities technological advances provide, recent studies indicated that secondary mathematics instruction still occurred mainly in a rote lecture format (White-Clark, DiCarlo, & Gilchriest, 2008). Computer and video games (henceforth referred to as digital games) are a ubiquitous presence in students' lives outside of school. By inviting digital games into the classroom, schools can potentially increase student learning and foster student motivation (Prensky, 2001). Digital games have become ingrained in the culture of today's learner, and in many cases are more engaging, immersive, motivating, relevant, and frequently provide for better overall learning experiences than those that occur in school (Gee, 2003; Simpson, 2005).

While some research findings have supported the use of digital games to improve student learning, limited empirical evidence exists pertaining to the relationship between digital games and student achievement, motivation for learning, or attitudes toward specific subjects of study (Girard, Ecalle, & Magnan, 2012; Harris, 2001; Hays, 2005; Sitzmann, 2011; Vogel, Vogel, Cannon-Bowers, Bowers, Muse, & Wright, 2006). Therefore, this study will analyze pre-existing data collected during a school program evaluation to determine the effects of playing a digital game called *Lure of the Labyrinth* (*Labyrinth*) on middle school students' mathematical achievement, motivation for learning mathematics, and attitudes toward mathematics.

Purpose of Study

The purpose of this study is to investigate data collected during a program evaluation of the effects of playing an immersive, serious digital game, *Lure of the Labyrinth,* on middle school students' mathematics achievement, motivation to be successful in mathematics, and attitudes toward mathematics.

Research Questions

The following questions guided this study:

1. What effects did playing a digital game (*Lure of the Labyrinth*) have on middle school students' mathematical achievement, as measured by internal assessments linked to the National Council of Teachers of Mathematics (NCTM) standards and benchmarks?
2. Did the effects of playing a digital game (*Lure of the Labyrinth*) on middle school students' mathematical achievement, as measured by internal assessments linked to NCTM standards and benchmarks, differ for boys versus girls?
3. What effects did playing a digital game (*Lure of the Labyrinth)* have on middle school students' motivation to learn mathematics, as measured by a modified version of the Course Interest Survey (CIS) (Keller, 1987)?
4. Did the effects of playing a digital game (*Lure of the Labyrinth*) on middle school students' motivation to learn mathematics, as measured by a modified version of the Course Interest Survey (CIS) (Keller, 1987), differ for boys versus girls?
5. Did playing a digital game (*Lure of the Labyrinth*) affect middle school students' attitudes toward mathematics as measured by the Fennema-Sherman Mathematics Attitudes Scale (FSMAS)?
6. Did the effects of playing a digital game (*Lure of the Labyrinth*) on middle school students' attitudes toward mathematics, as measured by the Fennema-Sherman Mathematics Attitudes Scale, differ for boys versus girls?

Definition of Variables

Sex. For purposes of this study, sex was an independent variable used to denote the biological category of male or female. *The Publication Manual of the American Psychological*

Association urges researchers to clearly delineate between socially constructed gender roles and biologically determined sexual categories (2010).

Playing an immersive, serious, digital game. For purposes of this study, *Lure of the Labyrinth* (*Labyrinth*) was the immersive, serious (definitions of these terms is provided below) digital game. *Labyrinth* is a Learning Games To Go product funded by a Star Schools US Department of Education grant and organized by Maryland Public Television. The Massachusetts Institute of Technology's Educational Arcade, Fable Vision, Maryland Public Television and ORC-Macro collaborated to develop *Labyrinth* ("Lure of the Labyrinth: About Us," n.d). Playing a digital game in this study refers to engaging with the game material and is the key independent variable. Students self reported the time spent playing *Labyrinth* in addition to the treatment group's allocated thirty minutes in school.

Attitudes toward mathematics. The Fennema-Sherman Math Attitudes Scales (FSMAS) measured learners' viewpoints regarding mathematics. Five subscales of the FSMAS were used (a) Attitude Toward Success Scale, (b) Confidence in Learning Scale, (c) Usefulness of Mathematics Scale, (d) Mathematics Anxiety Scale, and (e) Effectance Motivation Scale (see Appendix C).

Effectance motivation. White defined effectance motivation as the persistence in learning or exploring a topic or phenomenon for the sole reward of engaging in it (1959). Effectance motivation was measured by the FSMAS Effectance Motivation Scale, which was designed to measure effectance in mathematics as a dimension of motivation that ranged from lack of involvement to active enjoyment and seeking challenge in mathematics (Fennema & Sherman, 1976).

Mathematical achievement. In this study, mathematical achievement refers to the dependent variable of mathematical performance as measured by internal exams based on the National Council of Teachers of Mathematics (NCTM) standards and benchmarks for mathematics.

Motivation to learn mathematics. Motivation is an individual's situational specific desire to learn math course content and skills. Motivation to learn mathematics was a dependent variable measured by the Course Interest Survey and comprises four factors (a) attention, (b) relevance, (c) confidence, and (d) satisfaction (Keller, 2010) (Appendix B).

Definition of Terms

Attention, Relevance, Confidence, Satisfaction (ARCS) Model. An applied model of motivation that encompasses multiple motivation theories and is based on four factors: attention, relevance, confidence, and satisfaction (Keller, 1987). Course specific motivation can be measured within the ARCS model using the Course Interest Survey (CIS).

Digital games. In this study, digital games refers to products that provide digital information to one or more players, take input from the players and process that information according to a set of programmed game rules to alter the digital information provided to the players. Digital games can be played on multiple platforms including computers, hand-held devices (e.g. mobile phones, Gameboys, etc.) and consoles (e.g. Playstations, X-boxes, etc.) (Kirriemuir & McFarlane, 2004).

Fennema-Sherman Math Attitudes Scales (FSMAS). The FSMAS consists of nine individual scales, with each scale being composed of twelve statements about mathematics (six positive statements and six negative statements) that assess students' attitudes towards mathematics.

Immersive games. Games that affect players' senses through elements such as audio, graphics, and narrative so that they are drawn into the game world and the game interface ceases to be noticed by the player (Sweetser & Wyeth, 2005).

National Council of Teachers of Mathematics (NCTM). NCTM is the world's largest organization concerned with mathematics education. NCTM has published a series of math standards outlining a vision for school mathematics in the USA and Canada ("About NCTM: Mission, Vision, and Priorities," n.d.).

Serious games. In this study, serious games are immersive digital games based on constructivist learning theory to support specific learning goals (Annetta, Murray, Laird, Bohr & Park, 2006).

Significance of Study

The results of this evaluation study can benefit (a) educational leaders from a school improvement perspective; (b) teachers from an implementation perspective; and (c) game designers from a design perspective. New technologies and media are irrevocably changing the requirements of effective educational leadership in the twenty-first century (Collins & Halverson, 2009; Halverson, Grigg, Prichett, & Thomas, 2005; McLeod, 2011). Educational leaders must positively manage the infusion of new technologies into the educational environment, effectively support digitally engaged students (McLeod, 2011; Prensky, 2001), and prepare learners for technologically-rich contexts (Friedman, 2005; McLeod, 2011). Educational leaders are expected to make instructional decisions based on data. However, Groff and Mouza stated that lack of research on the efficacy of technology-based instruction poses a barrier to greater integration (2008). Ultimately, understanding whether digital games might be used to improve educational outcomes will inform educational leaders in making important decisions about access and resource allocation (Klopfer, et al., 2009).

Additionally, as teachers seek to integrate technology into their classrooms, they will require information regarding efficacy and best-practices for implementation (Ke & Grabowski, 2007). This study may provide valuable information for educators regarding the efficacy of using digital games to improve mathematical instruction.

While digital games and schools have followed different developmental paths, both comprise designed learning environments that encompass rich potential to enhance each other (Halverson, 2005). As game designers gain a clearer understanding of games' effects on learners, they will be better prepared to create more meaningful digital games (Jones & Kalinowski, 2007).

Organization of the Study

This dissertation is organized into five chapters. Chapter one introduces the study with a statement of the problem, research questions, and the significance of the study. Chapter two provides a literature-based background for the study by presenting research related to digital games and an overview of studies pertaining to mathematics. Chapter three presents the method of the study, which includes research design, instrumentation, and processes used for data collection and analysis. Chapter four presents the results of the study. Chapter five provides the discussion and conclusion for the study.

CHAPTER II

Review of the Literature

The literature review is presented in two sections, digital gaming and mathematics instruction, to present relevant information related to each of the major components within the study. The first section provides an overview of the use of digital games in education, implications of gaming for educational leaders, research on potential negative effects of gaming, and differences between males and females related to digital games. In addition, the first section examines pertinent research related to the effects of digital games on student achievement and motivation. The second section provides an overview of mathematics instruction and sex differences in mathematics. Further, the second section examines middle school students' mathematical achievement and attitudes toward mathematics.

Digital Games for Learning

The current generation of students' familiarity and comfort with Information Communication Technology (ICT) has led some researchers to refer to them as "digital natives" (Prensky, 2001), the "Net generation" (Tapscott, 1998), or the "millennials" (Howe & Strauss, 2000). Some writers have suggested that immersion in a technology-rich culture influences the skills and interests of this generation of students in ways that have significant implications for education (Annetta, Murray, Laird, Bohr, & Park, 2006; Frand, 2000; Gee, 2005; Klopfer, et al., 2009; Oblinger, 2004; Prensky, 2001; Tapscott, 1999). Rapid advances in ICT have led researcher to explore the possibilities of using digital games to foster learning based on the connection between play and learning (Koster, 2004; Oblinger, 2004; Shaffer, Squire, Halverson, & Gee, 2005).

Play and learning. “Education must begin in play, and the play spirit must pervade all work” (Hall, 1912, p. 112). The role of fun and play in learning has a long and distinguished history. In 1890, while writing about reasoning, James posited, “Intense interest or concentrated passion makes us think so much more truly and profoundly” (1890/2007, p. 367). Twenty years later, Dewey suggested that the ideal state for learning is “playful and serious at the same time” (1910/1997, p. 218). Piaget and Vygotsky, who laid the framework for constructivist learning theory, both asserted the importance of play as a primary mechanism for learning (Rieber, 1996).

According to Piaget, children’s play leads to cognitive development by providing opportunities for assimilation and accommodation (Ginn, 1995). Assimilation takes place when individuals attempt to integrate new information directly into existing mental structures or “schemata,” the mental representation of an associated set of perceptions, ideas, and/or actions. On the other hand, when individuals modify or change existing schemata to fit new information, accommodation occurs. From Piaget’s perspective, cognitive disequilibrium occurs when children encounter new information (Ginn, 1995). Primarily through various types of play, children resolve their cognitive dissonance either by assimilating the new information or changing their existing schemata to accommodate it (Ginn, 1995; Rieber, 1996). From a Piagetian perspective, the ongoing cycle of new inputs, dissonance and resolution is the learning process (Ginn, 1995).

Like Piaget, Vygotsky contended that play is pivotal in development (1978). Vygotsky’s social constructivist learning theory differentiates between two levels of development: actual developmental level as determined by independent problem solving, and potential developmental level as determined through problem solving with the guidance of more capable people (Ritterfeld & Weber, 2006). Vygotsky defined the zone of proximal development (ZPD) as the distance between these two levels of performance. According to Vygotsky, maximum learning occurs within an individual's ZPD and the interaction of play "creates the ZPD of the child. In play, a child always behaves beyond his average age, above his daily behavior; in play it is as though he were a head taller than himself" (Vygotsky, 1978, p. 102).

Van Eck asserted that digital games could be a powerful tool in increasing student learning by promoting curiosity and challenging students' abilities, which facilitate the Piagetian processes of cognitive disequilibrium and resolution (2006). Similarly, Ritterfeld and Weber argued that some digital games promote student learning by creating cognitive dissonance within the learner's ZPD (2006). Thus, digital games that challenge players with an attainable goal and allow players to collaborate with other people for assistance will have the greatest effect on student learning (Ritterfeld & Weber, 2006).

Researchers working with the *Quest Atlantis* digital game project, which is based upon a Vygotskian perspective of learning, play, and development, found that students who participated in the game world demonstrated significant increases in conceptual understanding on posttests as compared to pretests in the associated content areas (Barab, Dodge, Tuzun, Job-Sluder, Jackson, Arici, et al., 2007). In essence, digital games based on constructivist pedagogy, in which a goal is to create cognitive disequilibrium without exceeding the capacity of the player to succeed, can promote fun, play, and learning in a way that children find engaging (Ritterfeld & Weber, 2006).

Serious digital games. Advocates for the use of digital games in education stated that digital games are effective learning tools because they (a) use action instead of explanation, (b) create personal motivation and satisfaction, (c) accommodate multiple learning styles and skills, (d) reinforce mastery skills, and (e) provide interactive and decision making contexts (Charles & McAlister, 2004; Holland, Jenkins, & Squire, 2002).

Beginning in 2002, digital game designers joined with educators, the US military, corporations, and medical professionals to explicitly develop “serious games,” i.e. digital games to support specific learning goals (Annetta, Murray, Laird, Bohr & Park, 2006). The serious game movement embraces the power of immersive games to attract, engage, connect, and teach game players critical content in the games’ respective focus areas (Zyda, 2005). Characteristics of immersive digital games include a system of rewards and goals, which motivate players; a narrative context, which situates activity and establishes rules of engagement; learning content that is relevant to the narrative plot; and interactive cues that prompt learning and provide feedback (Dondlinger, 2007). Though they do not agree upon a single definition of serious games, experts within the movement concur that serious games include a relevant narrative, are engaging to play, are designed based on sound pedagogy, and have a learning goal (Michael & Chen, 2006). The serious game movement attempted to develop and distribute fun, pedagogically-sound digital games that had a learning goal and created a state of flow in players (Ellis, Heppell, Kirriemuir, Krotoski, & McFarlane, 2006). Flow describes a state of complete absorption or optimal experience that results from complete immersion in a goal driven activity (Csikszentmihalyi, 1996). Research studies have found that students whose lessons included serious games demonstrated stronger engagement in curriculum content and deeper understanding of concepts than those who did not use games (Barab, Dodge, Tuzun, Job-Sluder, Jackson, Arici, et al., 2007; Ellis, et al., 2006; Zyda, 2005). When comparing serious games with non-educational computer games, Zyda argued that the addition of sound pedagogy causes games to become serious and effective learning tools (2005).

Several researchers involved with serious games stressed that pedagogy must be subordinate to story and that the entertainment component must come first (Zyda, 2005).

Similarly, Prensky argued that effective digital games should be fun first and then should encourage learning (2001). Others have argued that in serious games, fun should be subordinate to learning goals (Michael & Chen, 2006). Koster suggested that fun and learning occur in tandem as he defined fun as “feedback the brain gives us when we are absorbing patterns for learning purposes” (2005, p. 96). Barab, Arici, and Jackson posited that helping, playing, and learning were intertwined activities that promoted student engagement and motivation (2005). Regardless of the debate over primacy and exact definitions, proponents of serious games agree that these games promote student motivation, which positively affects learning (Dondlinger, 2007; Ellis, et al., 2005).

MIT’s Educational Arcade is at the forefront of developing serious games based on constructivist learning pedagogy, flow, and engaging play. *Labyrinth* is a digital game developed by the Educational Arcade in conjunction with Maryland Public Television to support middle school students’ understanding of pre-algebra topics. Developers aligned *Labyrinth* with the Maryland voluntary state curriculum mathematics standards and benchmarks for grades six through eight (see Appendix A for details related to standards and benchmarks). *Labyrinth* is designed like a graphic novel with a strong storyline and embedded with math concepts, skill practice, and logic (see Figure 1). In order to save a beloved pet and ultimately the world, players disguised as monsters travel through the Labyrinth solving imbedded mathematical puzzles based on numeracy, ratios, proportions, and fractions to progress. For example, in the “Employee Cafeteria” (see Figure 2) the player must use ratios and proportions to solve the puzzle and serve the right amount of different foods to various monsters.

Figure 1. Avatar screenshot from *Lure of the Labyrinth*. This game page, from after a player chooses a monster disguise, shows the graphic novel style used throughout the game. Retrieved from http://labyrinth.thinkport.org/www/

Figure 2. Employee cafeteria puzzle screenshot from *Lure of the Labyrinth*. The puzzle shown in this figure furthers the game narrative and requires proportional thinking to solve. Retrieved from http://labyrinth.thinkport.org/www/

This study analyzed data collected during a program evaluation of using *Labyrinth* with middle school (grades 6, 7, and 8) mathematics students. The goal of this analysis was to determine whether\ mathematical academic performance, attitudes toward mathematics, and motivation for learning mathematics were positively affected during the program evaluation of playing *Labyrinth*. While not a requirement of playing *Labyrinth*, this program evaluation used the “Tasti Pet Communicator” function (see Figure 3), which allowed players to digitally collaborate during game play, but not during puzzles. A teacher monitored “Tasti Pet Communication” via the *Labyrinth* administration tools to ensure appropriate use.

Figure 3. Tasti Pet Communicator (TPC) screenshot from *Lure of the Labyrinth.* The TPC allows players to collaborate electronically while in gameplay mode. Retrieved from http://labyrinth.thinkport.org/www/

Student achievement and digital games. Over the past decade, educators who use digital games as instructional tools have reported positive effects on student learning in a variety of disciplines (Franklin, Peat, & Lewis, 2003; Koether, 2003; Lauer, 2003). The findings of empirical studies revealed that instructional digital games could promote learners' mathematics performance (Ke & Grabowski, 2007; Kebritchi, Hirumi, & Bai, 2010; Lee, Luchini, Michael, Norris, & Soloway, 2004; Pareto, Arvemo, Dahl, Haake, & Gulz, 2011; Pilli & Aksu, 2012; Shaffer, 1997).

Whereas several studies found a positive relationship between digital game use and improved student outcomes, large-scale reviews of studies of the effects of digital games revealed inconsistencies in results. Several published literature reviews provided an overview of the effects of digital gaming by analyzing, comparing and integrating the findings of reported research in order to draw conclusions about the effects of digital gaming. For example, a review conducted by Harris in 2001 found no clear causal relationship between academic performance and the use of digital games. Mitchell and Savill-Smith's review of the effects of digital gaming on learners' achievement was inconclusive and the authors suggested that significant research is needed to understand and identify any relationships (2004). Hays' 2005 meta analysis of 48 empirical studies found mixed results leading him to conclude that no evidence existed to indicate a positive correlation between instructional games and improved student achievement. However, a more recent review of 32 quantitative studies asserted that immersive digital games were more influential than traditional classroom instruction on learners' cognitive gains (Vogel et al., 2006).

Sitzmann (2011) performed a meta-analysis of 55 research reports relating to the instructional effectiveness of simulation games. She defined simulation games as "instruction delivered via personal computer that immerses trainees in a decision-making exercise in an artificial environment in order to learn the consequences of their decisions (Sitzmann 2011, p. 490). Sitzmann's analysis found that "declarative knowledge was 11% higher, procedural knowledge was 14% higher, and retention was 9% higher for trainees taught with simulation games, relative to a comparison group" (2011, p.489). She concluded that simulation games have the potential to enhance learning, but the "characteristics of simulation games and the instructional context were instrumental in determining the amount that trainees learned from simulation games relative to a comparison group" (Sitzmann 2011, p. 520).

Girard, Ecalle, and Magnan's (2012) meta-analytic review included only randomized controlled studies between 2007 and 2011 that used video games or serious games as learning tools. These authors defined serious games based on part of Marsh's definition "Serious games are digital games, simulations, virtual environments and mixed reality/media that provide opportunities to engage in activities through responsive narrative/story, gameplay or encounters to inform, influence, for well-being, and/or experience to convey meaning" (Marsh 2011, p. 63). Girard et al. (unlike Marsh) define video games as being the same as serious games but designed for entertainment rather than usefulness. Of the initial 30 experimental studies initially identified by Girard et al., all but eleven were excluded due either to issues with research design (lack of control group, random assignment to group, or pre- and post-test measures) or lack of focus on video games or serious games. Based on the eleven studies reviewed, two of the six classified as using serious games showed positive effects on learning and one had mixed effects. Out of the five studies classified as using video games, one showed positive effects on learning. Girard et al. concluded that insufficient evidence existed to determine whether digital games positively affect learning or whether serious games were more effective than video games at improving learning.

Harris (2001), Hays (2005), Vogel et al. (2006), and Girard et al (2012) noted that empirical research on the effectiveness of instructional games was difficult to assess due to the flawed methodology of several studies and the inconsistent definitions of key terms. These authors noted flaws in methodology including the lack of a comparison group, confounding variables such as instructional time differences between groups, and small sample sizes. Whereas many books and articles have discussed the topic of digital gaming in education, few empirical studies have addressed its effectiveness as a learning tool (Hays, 2005).

Some researchers have posited that the inconsistent results regarding the effectiveness of digital gaming on cognitive achievement is due in part to inherent differences in the types of games studied (Denis & Jouvelot, 2005). For example, researchers have compared the cognitive changes related to skill and drill educational games versus immersive digital games (Williams, 2005). Skill and drill games such as *Alge-Blaster*, which includes a tutorial of terms, systematic worked-out sample problems, and a graphing simulator, provide opportunities for students to practice already taught concepts. Educators consider skill and drill games to be easy to incorporate into their existing curricula and classroom practices (Squire, 2003). According to Denis and Jouvelot (2005), the didactical and linear progression of skill and drill digital games simply requires users to practice repetitive skills or rehearse memorized facts. Thus, skill and drill games fail in “transmitting non trivial (or previously assimilated) knowledge, calling again and again the same action patterns and not throwing the learning curve into relief” (Denis & Jouvelot, 2005, p. 464). In contrast, immersive games such as *Revolution,* a multiplayer game based on the American Revolution in which students can choose to assume the role of one of seven different social perspectives in 1775 Williamsburg, result in different game outcomes and are multilayered in order to intellectually and emotionally immerse players into the game world (Oblinger, 2004). Squire and Jenkins contended that immersive games “are not simply problems or puzzles; they are microworlds, and in such environments students develop a much firmer sense of how specific social processes and practices are interwoven, and how different bodies of knowledge relate to each other” (2003, p. 15). While studies have found evidence of improved academic achievement with both types of digital games, drill and practice games appear to have less of an effect on student learning than immersive digital games (Gee, 2005). Variations in student achievement outcome, such as those seen based on types of digital games, may be due in part to differences in the mediating factors, such as motivation, between digital games and learning (Williams, 2005).

Some studies have explored the mechanisms through which digital games have led to improved student achievement. Researchers have found that playing digital games encourages the brain to work more efficiently and thus take in more cognitive material than it would in a traditional learning setting (Pange, 2003; Perry & Ballou, 1997). Other studies found that playing digital games increased students’ attention (Yip & Kwan, 2006) and knowledge transfer (Shaffer et al., 2005). In addition, using digital games to support understanding resulted in dynamic (Rosas et al., 2003) and collaborative (Squire, Giovanetto, Devane, & Durga, 2005) learning environments, which positively affected learning.

Motivation and digital games. Engagement and motivation are difficult concepts to define and quantify (Gee, 2005). However, Csikszentmihalyi’s work on optimal experience theory or flow (1996; Shernoff, Csikszentmihalyi, Schneider, & Shernoff, 2003) has provided insight for digital game designers interested in developing engaging learning games (Gee, 2005;

Kiili, 2005). Kiili's research concluded that when used as instructional tools, the digital games that were associated with higher levels of student-perceived flow resulted in larger increases in student learning (Kiili, 2005). Flow, fun, and play within a pedagogically sound framework may all contribute to student motivation to sustain engagement in activities that promote learning.

Pink concurred that Csikszentmihalyi's "flow" state is a critical aspect of motivation for the 21st century (2009). Pink asserted that research has shown that autonomy, the desire to direct our own lives; mastery, the urge to continually improve at something that matters; and purpose, the yearning to do what we do in the service of something larger than ourselves are the key components of motivation. When people are in a flow state, they move toward mastery; however, one can never fully achieve mastery. Further, Pink asserted that children inherently possess a mastery mindset that leads them to frequently experience a state of flow. "Children careen from one flow moment to another animated by a sense of joy, equipped with a mindset of possibility, and working with the dedication of a West Point cadet. They use their brains and their bodies to probe and draw feedback from the environment in an endless pursuit of mastery" (p.130). Pink's analysis suggested that the current systems of providing external rewards and punishments to increase motivation undermines a mastery mindset, limits autonomy, and interferes with one's ability to achieve a flow state: "…introducing an 'if-then' reward to help develop mastery usually backfires. Thus schoolchildren who are paid to solve problems typically choose easier problems and therefore learn less. The short-term prize crowds out the long-term learning" (p. 58). In his blog, Pink referred to Gee's work and suggested that "Games, not grades!" may represent a way for educators to more effectively motivate student learning (2009). "When we think of games, we think of fun. When we think of learning we think of work. Games show us this is wrong. They trigger deep learning that is itself part and parcel of the fun" (Gee, 2005, p.15).

According to game theory, digital games support learning by being highly engaging, motivating, and interactive (Ju & Wagner, 1997; Kafai, 2001; Rieber, 1996). Several researchers found that playing digital games increased student motivation and interest levels, which led students to learn more quickly and thoroughly the material presented (Lardinois, 1989; Rieber, 1996; Romme, 2003; Rosas et al., 2003). Although studies have explored motivation in digital games, not all researchers agreed on the source of this motivation. Some attributed the compelling nature of games to their narrative context (Dickey, 2005, 2006; Waraich, Sharman, & Mitchell, 2004). Others found motivation was linked to goals and rewards within the games themselves or intrinsic to the act of playing (Amory, Naicker, Vincent, & Adams, 1999; Denis & Jouvelot, 2005). Denis and Jouvelot stated, "Intrinsic motivation pushes us to act freely, on our own, for the sake of it; extrinsic motivation pulls us to act due to factors that are external to the activity itself, like reward or threat" (2005, p. 462). These authors argued that competence, autonomy, and relatedness are factors that affect motivation. "Motivation also leads to the activation of efficient cognitive strategies for long-term memory issues like monitoring, elaborating or organizing information" (p. 463).

According to the Hierarchical Model of Motivation (Vallerand & Ratelle, 2002), motivation exists at three levels of generality. At the most global level, individuals tend to have a dispositional motivational orientation. However, at the most specific level, people adopt a particular motivational state in particular situations. Keller's (1987) ARCS model provides a way for conceptualizing motivation in a particular situation based on four factors (a) attention, (b) relevance, (c) confidence, and (d) satisfaction. See Figure 4 for description of each category and tactics for reinforcement.

The attention category in the ARCS model refers to the ability to capture the interest of learners, to pique their curiosity to learn, and to sustain their attention (Keller, 1992). Keller stated that students' attention in a learning environment can be reinforced through (a) perceptual arousal, (b) inquiry arousal, and (c) variability (Keller, 1987a). Methods of gaining attention through perceptual arousal include providing visually appealing graphics or unexpected actions to capture learner interest. Learners' initial perceptual attention can be maintained for greater lengths of time through inquiry arousal such as asking challenging or stimulating questions and integrating inductive and problem-solving approaches. Variability is important to prevent learner boredom due to too little variation or frustration from too much variation (Keller, 1992).

Keller's ARCS relevance factor refers to the degree that the learning experience is meaningful to students. Keller identifies three categories of tactics for enhancing relevance: (a) familiarity, (b) goal orientation, and (c) motivation matching (1992). Connecting learners' experiences to instruction increases familiarity. Ensuring that learners perceive the connections between what they need to know and the learning opportunities presented fosters goal orientation. Through goal orientation, the instruction is related to learners' goals. Encouraging learners to visualize achieving a goal or appealing to personal interests and learning styles are tactics related to motivation matching.

Within the ARCS model, the confidence category encompasses the basic constructs of Expectancy Theory in which learners' motivation is based on their evaluation of self-efficacy for success, perceptions of whether the effort will lead to desired outcomes, and the value of the expected outcomes (Vroom, 1964). Keller (1992) presented three methods of instilling confidence in learners: (a) learning requirements, (b) success opportunities, and (c) personal responsibility. Providing learners with clear learning objectives and expected outcomes of the instruction helps to clarify the learning requirements. Success opportunities occur when learners are challenged at an appropriate level. Personal responsibility involves ensuring that learners feel that they succeeded due to their ability rather than because the task was too easy, they were lucky, or other external factors (Keller, 1992).

Satisfaction within the ARCS model refers to learners' positive feelings about their learning experiences, including affirmation that the instructional content was relevant and that they had the ability to learn the material. Keller identified three kinds of tactics to improve learner satisfaction: (a) natural consequences, (b) positive consequences, and (c) equity. Natural consequences refer to intrinsic reinforcement or an internal desire to learn and can be supported by providing opportunities to use newly acquired knowledge in a real or simulated setting. Through constructive feedback, learners experience positive consequences or extrinsic rewards in which they are recognized for their accomplishments, either verbally or through actual rewards. Equity includes learner perceptions of fair and consistent treatment, which are essential to motivation regardless of the context (1987).

Category	Description	Tactics for Reinforcement
Attention	The ability to capture interest, pique curiosity to learn, and sustain attention	**Perceptual arousal**: novel and surprising activities **Inquiry arousal:** posing questions and problems **Variability**: too little variation can lead to boredom, too much can lead to frustration
Relevance	The degree that the learning experience is meaningful to students	**Familiarity**: adapting instruction according to learners' background knowledge **Goal orientation**: presenting instructional goal **Motivation matching**: providing instructional strategies based on learners' profiles
Confidence	A combination of learners' evaluation of self-efficacy for success, perceptions of whether the effort will lead to desired outcomes, and the value of the expected outcomes	**Learning requirements**: clear learning objectives and expected outcomes of instruction **Success opportunities**: learners are challenged at an appropriate level **Personal responsibility**: learners perceive success is due to their ability and effort
Satisfaction	Students' positive feelings about their learning experiences	**Natural consequences**: intrinsic reinforcement supported by providing opportunities to use newly acquired knowledge in a real or simulated setting **Positive consequences:** constructive feedback or extrinsic rewards for accomplishments **Equity**: perceptions of fair and consistent treatment

Figure 4. ARCS Model. This figure shows each of the four ARCS categories, including descriptions and recommended tactics for reinforcement.

Researchers have used the ARCS model to evaluate motivations of learners in a variety of learning settings including E-learning environments and web-based distance settings (Keller & Suzuki, 2004), hypermedia contexts (Carson, 2006), and classroom face-to-face setting

(Small, Zakaria & El-Figuigui, 2004). In this study, the effects of *Labyrinth* on learners' motivation were measured by using the Course Interest Survey (see Appendix B), which is based on the ARCS model of motivation.

Sex differences and digital games. According to research conducted by Pew Internet and American Life Project in 2008, ninety-nine percent of teen boys and ninety-four percent of teen girls in the United States reported playing digital games (Lenhart et al., 2008). Traditionally, society has perceived digital games as a male-dominated domain (Cassell & Jenkins, 1998). Cassell and Jenkins's work *From Barbie to Mortal Kombat: Gender and Computer Games* pointed out the variety of ways digital games reflected gender stereotypes, relied on male-dominated central characters, and focused on stereotypically male areas of interest, such as fighting and racing (1998). Perhaps in part due to these issues, research has found that males not only play games more often, but they also play different types of games (Bonanno & Kommers, 2005; De Jean, Upitis, Koch, & Young, 1999; McFarlane, Sparrowhawk, & Heald, 2002). Further, males and females appeared to have significantly different attitudes toward the use of video games (Bonanno & Kommers, 2008; Bourgonjon, Valcke, Soetaert, & Schellens, 2010).

Bonanno and Kommers explored differences between males and females in the digital gaming practices of 367 European students aged 16-18 (2005). Based on students' survey responses, Bonanno and Kommers found that males spent significantly more time engaged with digital games ($p < .001$) with an average of 6.7119 hours per week as compared to just 2.4917 per week for females. In addition, Bonanno and Kommers found a significant difference ($p < .001$) in the types of digital games preferred by males (role playing games) versus females (puzzle games). Lastly, these researchers found that males exhibited greater diversity in types of games preferred. Further work by Bonanno and Kommers explored sex differences in the affective component, perceived usefulness, perceived control, and behavioral components of digital gaming in a learning context reported by 170 Maltese college students aged 16-18 (2008). Neither males nor females reported overall negative affect related to playing games, but males reported less hesitancy and greater confidence in game playing than females ($p < .015$). In terms of usefulness, both males and females viewed gaming as an "interesting and imaginative" way to learn and as an "efficient and effective" learning experience, but females were more skeptical of games' learning potential than males ($p < .076$). Regarding perceived control, males reported greater confidence in their ability to solve within-game problems ($p < .002$) as well as to command the gaming device itself ($p < .002$). Further, females were significantly more likely to express a need for "guidance and support from a more competent person" to effectively play ($p < .002$). Related to gaming behavior, females reported a significantly higher level of avoidance of game playing outside of a learning context ($p < .001$). However, both sexes reported that they did not avoid using digital games for learning.

Like Bonanno and Kommers (2008), Bourgonjon et al. (2010) explored attitudes related to digital games. Bourgonjon et al. surveyed 858 Flemish secondary students ranging from ages 12 to 20. Bourgonjon et al. found that males spent significantly more time playing digital games ($p < .001$), with an average of 6.96 hours a week, than females, who averaged 2.16 hours per week. Based on an extension of Davis's technology acceptance model (1989), Bourgonjon et al. examined students' acceptance of digital gaming for learning in the classroom. They found that both males and females reported perceiving digital game use in the classroom as representing valued learning opportunities, but males expressed a preference for learning using digital games whereas females did not ($p < .05$). Pathway analysis supported the hypothesis that sex significantly influenced prior gaming experience ($p < .001$) and ease of use ($p < .001$), which as mediating factors accounted for 63% of the sex variance in preference for learning with digital games (2010).

Sex differences in digital game use appear to be deeply rooted and persistent, leading some researchers to express concern that digital games might alienate female students (Carr & Pelletier, 2008; Dawes & Dumbleton, 2002). However, Ke (2008) and Papastergiou (2009) found that the hypothesized impact of sex appeared to dissipate during the implementation phase of working with immersive learning digital games and ultimately digital games were equally effective and motivating for both male and female students.

Potential negative effects of digital gaming. The two concerns most often raised related to digital games are the potential for gaming to lead to increased aggression and “addictive-like” gaming behavior leading to negative psychosocial outcomes (Gentile et al., 2011). To avoid these issues, the original program evaluation focused on a game that had no violent content and limited playing time.

Overall, the research on the link between playing violent digital games and aggression is equivocal. Some research reported a relationship between violent video games and increased aggression (Anderson & Dill, 2000; Bartholow & Anderson, 2002; Bartholow, Bushman, & Sestir, 2006). In one of the most commonly cited studies, Anderson and Dill stated that their research with 210 undergraduate students showed a causal link between playing violent digital games and increased aggression (2000). However, Ferguson et al. pointed out that Anderson and Dill only reported statistical significance on one of the four measures of aggression used and that this measure would not have been significant had Anderson and Dill used a statistical correction (e.g. Bonferroni correction) for the multiple comparisons in their study (2008). Lastly, Ferguson raised concerns that the noise blast instrument used by Anderson and Dill for measuring aggression had not been validated (2007).

Other researchers reported no evidence of a predictive link between violent video games and aggression (Ferguson et al., 2008; Weigman & Van Schie, 1998; Williams & Skoric, 2005). For example, Ferguson et al. performed two studies that undermined the construct of a direct link between violent video games and aggression (2008). In the first study Ferguson et al. followed a procedure similar to Anderson and Dill's (2000), but used a standardized measure of the noise blast system for the aggression variable. Their analysis revealed no significant increase in aggression among the participants who played the violent video game versus the non-violent game. In the second study, Ferguson et al. used several self-reporting instruments with 428 university students. Their results indicated that exposure to family violence (verbal and/or physical) was a predictor of trait aggression and violent criminal acts. When exposure to family violence was controlled for, the researchers found no link between violent video games and criminal acts or trait aggression (2008).

Meta-analyses of studies that investigated the link between violent video games and aggression have also produced mixed results. Anderson and Bushman (2001) and Anderson (2004) found a small but significant positive relationship between violent games and aggression, however three other meta-analyses found no link (Ferguson, 2007; Sherry, 2001; Sherry, 2007). Given the potential link between violent games and aggression, the program evaluation that generated the data for this study focused on a digital game that does not contain violent content.

Although researchers continue to debate whether video game over-use constitutes addiction, the Diagnostic and Statistical Manual of Mental Disorders has yet to address the topic (Holden, 2010). However, emerging research indicates that under certain circumstances video gaming is associated with dysfunctional behaviors that harm a person's social, occupational, family, school, health, and/or psychological functioning (Gentile et al., 2011). In the research literature, dysfunctional over-engagement in gaming is often referred to as pathological gaming (Gentile et al., 2011). Using screening instruments based on the constructs for pathological gambling, studies in the US, China, Taiwan, Australia, and Germany have found incident rates of pathological gaming to range from 7.5-11.9% of the general population (Gentile, 2009; Grüsser, Thalemann, & Griffith, 2007; Ko, Yen, Yen, Lin, & Yang 2007; Peng & Li, 2009; Porter, Starcevic, Berle, & Fenech, 2010). However, research on the correlation between over-use of gaming and negative psychology and social functioning have yet to determine directionality or causality (Gentile et al., 2011). In other words, it has not yet been determined whether issues such as depression led to pathological gaming, were a consequence of over-engagement in gaming, or if both excessive gaming and depression developed as a result of some other factor. Although research on pathological gaming is inconclusive, the program evaluation limited participants' exposure to digital gaming to control for any potential negative effect of over engagement.

Digital gaming and K-12 leadership. Halverson et al. contended that increased accountability has resulted in a revolution in educational leadership, in which "changing the organizational conditions for improvement across schools is the central task of school leaders" (Halverson, Grigg, Prichett, & Thomas, 2005, p. 3). Traditional roles in educational leadership focused on tasks such as defining the school mission, promoting quality instruction through teacher evaluation, maintaining high academic expectations, and developing a strong school culture (Collins & Halverson, 2009). In addition to these duties, today's school administrators cope with school performance measures, high stakes instructional standards, customer satisfaction demands, increasing staff development needs, and the rapid influx of computer technology. Perhaps the biggest challenge for the educational leader lies in the ability to break away from obsolete pedagogical practices within the school and replace them with current, research-based practices (Collins & Halverson, 2009; Halverson, 2005; Halverson, Grigg, Prichett, & Thomas, 2005).

This focus on improved pedagogical practices is represented in the International Society for Technology in Education (ISTE) technology standards for educational leaders. A quarter of the performance indicators for the five standards reflect the increased expectation of instructional leadership related to technology for administrators (The International Society for Technology in Education [ISTE], 2009). The ISTE performance indicators for educational leaders include the following:

- "Ensure instructional innovation focused on continuous improvement of digital-age learning" (ISTE, 2009, p.1).
- "Model and promote the frequent and effective use of technology for learning" (ISTE, 2009, p.1).
- "Ensure effective practice in the study of technology and its infusion across the curriculum" (ISTE, 2009, p.1).
- "Stay abreast of educational research and emerging trends regarding effective use of technology and encourage evaluation of new technologies for their potential to improve student learning" (ISTE, 2009, p.1).
- "Lead purposeful change to maximize the achievement of learning goals through the appropriate use of technology and media-rich resources" (ISTE, 2009, p.1).

The historical, traditional loose coupling of instructional practice and educational leadership "granted individual teachers space for instructional innovation" (Halverson, 2005, p. 3). However, today's climate of accountability requires a closer coupling in which school leaders fully take the reigns of instructional leadership (Collins & Halverson, 2009; Halverson, 2005; Halverson, Grigg, Prichett, & Thomas, 2005). In the age of data-driven school reform, the limited and ambiguous research base regarding the effectiveness of digital gaming as an instructional technique presents educational leaders with a barrier against greater adoption (Groff & Mouza, 2008). Theoretically and empirically, digital games show promise for improving student outcomes (Barab, et al., 2007; Gee, 2005; Prensky, 2006; Vogel et al, 2006), making it incumbent upon educational leaders to initiate, support, and review systematic investigations of gaming in education (Collins & Halverson, 2009; Halverson, 2005). "Thus, although the inspiration for instructional innovation often comes from teachers, the responsibility for realizing the power of game-based learning environments across schools lies mainly with school leaders" (Halverson, 2005, p. 3).

Middle School Mathematics

Middle school mathematical experiences set the stage for students' ongoing achievement in mathematics (Nathan & Koellner, 2007). In middle school, mathematical concepts become abstract and increasingly difficult (Leinhardt, Zaslavsky, & Stein, 1990; Utsumi & Mendes, 2000). In addition, "there is growing evidence that individual differences in academic performance cannot be explained as solely the result of differences in general ability but appear as the product of complex and dynamic interactions among cognitive, affective, and motivational variables" (Volet, 1997, p. 235).

Sex differences and mathematics. Over the past four decades, researchers have extensively explored the relationship between sex and mathematics performance. Findings in the 1970s and 1980s consistently showed that boys outperformed girls in most areas of mathematics (Benbow & Stanley, 1980; Maccoby & Jacklin, 1974). Some recent studies suggested that gender differences on mathematics achievement tests have disappeared in the K-12 arena (Freeman, 2004; Hyde, 2005; Hyde, Lindberg, Linn, Ellis, & Williams, 2008). Other researchers contended that gender differences in mathematics achievement remain significant, begin during the early elementary years, and are pronounced among high achieving students (Robinson & Lubienski, 2010; Stubits, Zackon, Roberts, Siegal & Flanagan, 2011).

Robinson and Lubienski analyzed the mathematical achievement of 7,075 students from fall of kindergarten through spring of grade eight, using data from the Early Childhood Longitudinal Study, Kindergarten Class of 1998-1999. Based on both t-scores and scaled scores, they found no statistical differences between the mean scores for males versus females at the kindergarten level. However, as students progressed through elementary, the mathematical achievement gender gap favoring boys increased to a high of 0.24 SDs ($p < .001$) for the standardized score and six points on the scaled score. The mean difference for boys' and girls' mathematical achievement decreased in middle school, but remained significant through spring of grade eight at a level of 0.12 SDs ($p < .001$) on standardized scores and a 2.5 point difference on scaled scores. In addition to analyzing mean scores, Robinson and Lubienski also performed a quantile regression of students' scores, which found that a smaller proportion of girls than boys performed at the higher levels, 75^{th} and 90^{th} percentile, of mathematical achievement (2010).

Stubits et al.'s 2011 analysis of longitudinal mathematical achievement data for five cohorts of students ($n = 554$) from kindergarten through grade four also found no significant gender difference at the kindergarten level, which indicated that both genders begin schooling on an equal mathematical level. Additionally, Stubits et al. investigated the pattern of significant achievement change (effect size of 0.8 or more) for boys and girls as they progressed from kindergarten through grade four. Their results indicated that girls were twice as likely to show mathematical achievement decline as compared to boys. In addition, boys were twice as likely to demonstrate significant achievement gains as compared to girls.

Robinson and Lubienski (2010) and Stubits et al. (2011) attributed some of the sex differences in mathematics achievement to girls having greater societal pressure to conform to rules, and thus to do math the way the way they were taught. By contrast, these researchers posited that boys have greater freedom to creatively approach mathematics, which results in greater mathematical achievement on more abstract and complex problems (Robinson & Lubienski, 2010; Stubits et al., 2011). While all students are likely to benefit from opportunities to creatively experiment with mathematical problem solving (Van de Walle, 2010), encouraging girls to engage in creative approaches to mathematics may help reduce the sex mathematical achievement gap (Stubits et al., 2011).

Regardless of achievement levels, girls' self-perceptions of mathematical competence continued to be lower than that of boys (Eccles, Wigfield, Harold, & Blumenfeld, 1993; Else-Quest, Hyde, & Linn, 2010; Ginsburg, Cooke, Leinwand, Noell, & Pollock, 2005; Pajares, 2005; Wigfield, Eccles, Mac Iver, Reuman, & Midgley, 1991; Wigfield et al., 1997). The Trends in International Mathematics and Science Study (TIMSS) and Program for International Student Assessment (PISA) data showed sex differences in perceptions of mathematical competence across a wide range of participating countries (Else-Quest et al., 2010; Ginsburg et al., 2005). Additionally, girls in high school showed less interest in and enjoyment of mathematics as compared to boys (Halpern et al., 2007; NCES, 2004; NSF, 2011). Pajeras indicated the sex confidence gap is notable by middle school and firmly established by high school graduation (2005). Catsambis found sex differences in perceptions of competence continued into the university level (2005). Addressing potential sex differences in mathematics confidence remains important, as research indicates students' self-perceptions of ability and expectancies for success are among the strongest predictors of future performance and task choice (Eccles et al., 1989; Wigfield, 1994). In addition, sex differences in students' relationships with mathematics may be part of why fewer women are enrolled in math-related sciences such as engineering, computer and information science, physical science, and chemistry in US universities (Halpern et al., 2007; NCES, 2004; National Science Foundation [NSF], 2011).

Attitudes and mathematics. Several quantitative studies have found a positive relationship between attitudes toward mathematics and achievement in mathematics (Kadijevich, 2008; Ma & Kishor, 1997; Ma & Xu, 2004; Minato & Kamada, 1996; Singh, Granville, & Dika, 2002; Tocci & Engelhard, 1991). Ma and Kishor's meta-analytic review found an overall positive population correlation of 0.12 between attitudes and achievement in mathematics across studies (1997). Kadijevich's examination of The Trends in International Mathematics and Science Study (TIMMS) 2003 data found significant correlations between attitudes and mathematic achievement scores in 31 out of the 33 participating countries (2008). Research into the directionality or causality of the correlation between mathematical attitudes and achievement is in its infancy. However, most studies confirmed a strong relationship (Ma & Xu, 2004).

Furthermore, research indicated that the middle school years might be a critical time for students to develop their attitudes toward mathematics (Ma & Kishor, 1997; Ma & Xu, 2004; Middleton & Spanias, 1999; Utsumi & Mendes, 2000; Wolf & Blixt, 1981). Building and maintaining a positive attitude toward mathematics is not common; for example only about 7% of Americans reported positive school experiences with mathematics (Jackson & Leffingwell, 1999). Proponents of learner-centered educational strategies suggested that digital technologies, already integrated into most aspects of students' lives, could promote improved mathematical understanding, motivation, and positive attitudes toward the subject area (Green & Hannon, 2007).

Mathematical Achievement. Increasingly, schools measure student achievement in relation to an established set of standards and benchmarks. While there does not exist a world-wide agreed upon set of mathematics standards and benchmarks, the work of the National Council of Teachers of Mathematics (NCTM) played a prominent role in the development and refinement of many systems of standards and benchmarks (Van de Walle, Karp, & Bay-Williams, 2010).

In 2000, the NCTM published *Principles and Standards for School Mathematics,* which contained a comprehensive list of pre-kindergarten through grade 12 standards and benchmarks for mathematics (NCTM, 2000). *Principles and Standards* was based on five years of drafting and revising involving a wide variety of people including mathematicians, educators, researchers, and policymakers (Kilpatrick, 2003). *Principles and Standards* comprises three sections: (a) the six principles fundamental to high-quality mathematics education, (b) the five content standards which are common to all grades, and (c) the five process standards through which students should acquire and use mathematical knowledge (NCTM, 2000). The five content standards, (a) Number and Operations, (b) Algebra, (c) Geometry, (d) Measurement, and (e) Data Analysis and Probability, provided guidance to teachers and leaders on what mathematics content should be taught (NCTM, 2000). Despite differences across US states, countries, and international schools regarding mathematics standards and benchmarks, *Principles and Standards for School Mathematics* remains a key reference in developing and aligning mathematic curriculum (Reys, 2008).

In 2006, the NCTM released *Curriculum Focal Points for Prekindergarten through Grade 8 Mathematics: A Quest for Coherence* to serve as a guide for key mathematical learning targets (NCTM, 2006). Several US states including Maryland based their voluntary state curricula in mathematics on *Focal Points* (Achieve, 2009). *Labyrinth* is aligned with Maryland's voluntary state mathematical curriculum. While the use of standards and benchmarks in assessing student achievement in mathematics is debatable (Ellis, 2008), the NCTM *Principles* and *Focal Points* are important and respected references for identifying key learning targets in mathematics (Van de Walle et al., 2010).

Summary

In the developed world, digital games play a significant role in the lives of most of today's students, either through direct play or through having friends and family who are gamers (Chatfield, 2010; Simpson, 2005). However, schools largely continue to follow a 19th century approach to teaching, which does not take into account the cultural perspective and needs of today's "digital natives" (McLeod, 2011; Prensky, 2001; Prensky, 2006). The serious game movement is attempting to bridge the disconnect between formal learning opportunities and students' interest by creating games that are fun, are pedagogically sound, and have a specific learning goal (Ellis et al., 2006; Michael & Chen, 2006). Serious digital games may improve student mathematical achievement by affecting the mediating factors of attitudes toward mathematics and motivation (Ke, 2008; Ke & Grabowski, 2007). However, research indicates that boys and girls have different relationships with both mathematics (Robinson & Lubienski,

2010) and digital games (Bonanno & Kommers, 2008). Despite indicators of the potential of digital games to improve student achievement, there exist limited empirical evaluation studies with strong methodology (Harris, 2001; Hays, 2005; Vogel et al., 2006).

CHAPTER III

Methods

Purpose of the Study

The purpose of this study was to examine data collected during a program evaluation to explore the effects of playing a serious digital game, *Lure of the Labyrinth*, on middle school students' mathematics achievement, motivation to be successful in mathematics, and attitudes toward mathematics.

Research Questions

The following questions guided this study:

1. What effects did playing a digital game (*Lure of the Labyrinth*) have on middle school students' mathematical achievement, as measured by internal assessments linked to the National Council of Teachers of Mathematics (NCTM) standards and benchmarks?
2. Did the effects of playing a digital game (*Lure of the Labyrinth*) on middle school students' mathematical achievement, as measured by internal assessments linked to NCTM standards and benchmarks, differ for boys versus girls?
3. What effects did playing a digital game (*Lure of the Labyrinth)* have on middle school students' motivation to learn mathematics, as measured by a modified version of the Course Interest Survey (CIS) (Keller, 1987)?
4. Did the effects of playing a digital game (*Lure of the Labyrinth*) on middle school students' motivation to learn mathematics, as measured by a modified version of the Course Interest Survey (CIS) (Keller, 1987), differ for boys versus girls?

5. Did playing a digital game (*Lure of the Labyrinth*) affect middle school students' attitudes toward mathematics as measured by the Fennema-Sherman Mathematics Attitudes Scale (FSMAS)?
6. Did the effects of playing a digital game (*Lure of the Labyrinth*) on middle school students' attitudes toward mathematics, as measured by the Fennema-Sherman Mathematics Attitudes Scale, differ for boys versus girls?

Research Design

This study analyzed data collected during a program evaluation of the efficacy of incorporating digital games into the middle school math program of an international school. The program evaluation was conducted using a pretest posttest design. One-hundred sixty-eight middle school students (grades 6, 7, and 8) attending the International School of Kenya (ISK) during the 2011-12 academic year participated in a school sponsored program evaluation.

ISK was established in 1976 and occupies a 25-hectare campus in suburban Nairobi. ISK comprises an elementary (pre-kindergarten to grade 5), middle (grade 6 to grade 8), and high school (grade 9 to grade 12), which issues an American High School diploma and International Baccalaureate Diploma. The Council of International Schools (CIS) and the Middle States Association of Schools and Colleges (MSA) accredit ISK. In academic year 2011-12, ISK has approximately 850 students, representing over 75 nationalities. ISK employs 109 teachers and administrators who represent 14 nations of which approximately 60% are North American, 18% are European, and 16% are African.

The program evaluation employed stratified random sampling to ensure equal representation of grade levels and boys and girls. The middle school administrative assistant separated the189 ISK middle school students into six strata based on grade level (6,7, or 8) and sex (male or female). The middle school principal used a random number generator to assign 14 students to the treatment and control groups from each stratum. Since grade level was not a variable of interest but was used only to ensure equal representation in the samples, the stratified random assignment resulted in 84 students, 42 boys and 42 girls, in the control and treatment groups. The 21 remaining students were assigned to participate along with their peers so that in each grade-level the control and treatment groups were even, but their information was not included in the data set. Neither teachers nor students were aware of which students in each group were actual participants. This study analyzed data collected during the program evaluation related to three dependent variables (a) achievement, (b) motivation, and (c) attitude toward mathematics. Each of these variables was measured at the beginning and end of the program evaluation. Thus, change on each variable from before to after playing *Labyrinth* was analyzed.

ISK middle school students attended four math classes of 80 minutes and one math class of 65 minutes in a two-week cycle. All students received the same amount of classroom time devoted to math. During the program evaluation, the experimental group played *Labyrinth* for thirty minutes once a week for nine weeks while the control group engaged in more typical mathematical practice activities. Both groups focused on the same curricular benchmarks (see Appendix A for details on curricular standards and benchmarks). Both groups had access to a variety of materials, such as assigned problem sets, the National Library of Virtual Manipulatives, and textbook readings to practice and solidify concepts outside of class. The treatment group was able to use *Labyrinth* outside of class time, but was not encouraged or assigned to do so. The control group was not given access information to *Labyrinth*; however, *Labyrinth* is openly available without access codes. Treatment group participants were asked to log the time spent playing *Labyrinth* outside of class. Both groups were asked to self-report time spent playing *Labyrinth* outside of class on the posttest survey. It was not part of the program evaluation design that students play *Labyrinth* outside of the classroom, however this self-reported information may have provided insight or indicated directions for future exploration if group differences had not been found.

The program evaluation occurred while students in grades six, seven, and eight were studying proportions, ratios, and fractions. Students in the treatment groups were directed to play Wing One in *Labyrinth,* which is focused on the same mathematical content areas taught during the program evaluation. While students in the treatment groups were playing Labyrinth, students in the control group were practicing already taught material through completion of problem sets. Although collaboration was not a variable of interest in the study, collaborative problem solving occurred in both the treatment and control groups. In addition to regular informal forms of shared physical space collaboration that occurred in both groups, the treatment groups students were taught to use the Tasti Pet Communicator (TPC) which is *Labyrinth's* within game communication device (see Figure 3). *Labyrinth* was designed to align with the Maryland Voluntary State Curriculum (VSC) and is correlated with the National Council of Teachers of Mathematics (NCTM) standards and expectations for grades six through eight (Lure of the Labyrinth, n.d.). While, the Maryland VSC, NCTM and ISK standards and benchmarks are not exactly the same, they are similar. See Appendix A for a detailed comparison of the VSC, NCTM, and ISK standards covered in Wing One of Labyrinth and addressed in classroom instruction during the treatment phase of the program evaluation.

Internal validity. Inferences regarding cause and effect can be drawn due to the program evaluation use of an experimental design and random stratified assignment of participants. However, the treatment and comparison groups were not isolated from each other, thus social interaction could possibly threaten internal validity. To minimize potential threats to internal validity, parents, teachers, and administrators received training and support to understand the importance of preserving group membership and avoiding any attempts to equalize group experience. Two short sessions were held for teachers and administrators. The first one occurred before the program evaluation began and focused on the scientific rationale for random sampling and the importance of preserving the integrity of group membership. Potential responses to student questions about group assignment and experience were role-played and ISK faculty had an opportunity to raise questions or concerns. A second short faculty session was held at week three of the study to provide an additional opportunity to address any questions that occurred as well as to reinforce the importance of maintaining group membership. Parents attended one informational session before the program evaluation began which included the scientific rationale for random sampling, the importance of preserving the integrity of group membership, and examples of how to respond to their children's questions. Parents and teachers were encouraged to ask question as they arose during the program evaluation.

Population and Sample

This study analyzed data collected during a program evaluation. The population for the program evaluation was 168 middle school students who attended the International School of Kenya (ISK) in 2011-12. The math course sequence in ISK's middle school comprises Integrated Mathematics 1 (typically grade 6), Integrated Mathematics 2 (typically grade 7), and Math 1 (typically grade 8). Tracking does not occur in ISK's middle school. However, students whose math performance is significantly different from grade level may be placed in either a lower or higher course than their grade. In addition, ISK support services (i.e. study skills workshop) and external assistance (i.e. tutoring) are recommended or required for students whose mathematical achievement is not at grade level.

Stratified random group assignment resulted in 168 ISK middle school students participating in the program evaluation evenly distributed across the three grades with a balanced numbers of boys and girls at each grade level. Students in the three middle school grades ranged from 10-15 years old and represented over 40 nations; approximately 35% were North American, 34% were European, 16% were African, and 15% were from other regions. However, in international schools like ISK, nationality is not easily quantifiable and may not serve as an accurate descriptor as many students have strong affiliations to multiple countries and several are citizens of more than one country. While there is great diversity of color, native language, and nationality, ISK middle school students are from privileged backgrounds. They are the children of wealthy Kenyan families, foreign diplomats, foreign development workers, or foreign company employees. Due to their parents' work, ISK students have access to many of the privileges of wealth such as attending a private school like ISK and the presumption of continuing their education at universities around the world. ISK's middle school accepts students with documented learning disabilities and provides support through an inclusion model. However, ISK does not admit students whose learning differences cannot be met within a regular classroom setting with minimal modifications or who are performing more than two years below grade level. Approximately 10% of ISK middle school students have Individual Education Plans (IEPs) based on their special needs and an additional 10% receive extra learning support but do not meet the criteria for implementation of an IEP. ISK's middle school accepts students with low English language proficiency and provides them with English for Speakers of Other Languages and inclusion support in the classroom.

External validity. Based on Campbell's Proximal Similarity Model, it is likely that generalizations about the results of this study can be made to other groups of similar students, in a similar setting, in a similar place and time period (1986). Thus, the results from this study can be cautiously generalized to students in other international schools and private college preparatory institutions.

Instrumentation

The study has three dependent variables: mathematical achievement, motivation for learning, and attitudes toward mathematics. The following sections discuss the instruments that were used to measure each of these variables.

Mathematical achievement. Mathematical achievement was measured by ISK internal math assessments. Each ISK assessment is linked to curricular benchmarks that were derived from the National Council of Teachers of Mathematics' (NCTM) standards and benchmarks (see Appendix A for more details related to curriculum standards and benchmarks).

Construct validity and reliability. The NCTM's *Principles and Standards for School Mathematics* (2000) and *Curriculum Focal Points for Prekindergarten through Grade 8 Mathematics: A Quest for Coherence* (2006) are widely respected guidelines for mathematical instruction (Van de Walle, 2010) and provide a measure of content validity for ISK's mathematical assessments.

The Programme for International Student Assessment (PISA), which was developed through the Organisation for Economic Cooperation and Development (OECD), developed the mathematical constructs and framework used in the Mathematical Literacy component of the International School Assessment (Australian Council of Educational Research [ACER], n.d.a). A team of international mathematics experts developed the PISA definitions and framework for mathematics. Thus, the PISA constructs for mathematics have strong international academic endorsement, as does the International School Assessment's (ISA) Mathematical Literacy construct validity by extension. Measures of internal consistency of 40 ISA Mathematical Literacy instruments from 2002 through 2006 for grades three to ten ranged from 0.80-0.88 suggesting that each iteration of the instrument had solid reliability (ACER, n.d.b). A Pearson Correlation comparing a randomly selected group of 30 students' scores on internal ISK mathematics assessments and their scores of Mathematical Literacy on the International School Assessment (ISA) demonstrated a strong relationship (r = .82) which provides criterion validity for ISK assessments as a measure of mathematical achievement.

Motivation for learning. Student motivation related to math class was measured using a modified online version of the Course Interest Survey (CIS). The CIS was developed by Keller based on his Attention, Relevance, Confidence, and Satisfaction (ARCS) model of motivation (1987). The ARCS model suggested that motivation can be broken down into four categories (1) attention, (2) relevance, (3) confidence, and (4) satisfaction (Keller, 1987). Keller stated that the attention category refers to those things that gain the learner's attention, build curiosity, and sustain active engagement (Keller, 2008b). In Keller's model the relevance category refers to strategies and concepts that build connections between the instructional environment and the learner. The ARCS confidence category relates to the increase in motivation to learn that people experience when they perceive that they can succeed in mastering a task. Thus, the confidence category refers to variables related to self-efficacy, attribution theory, self-determination theory, and goal oriented theory. Keller's last category, satisfaction, pertains to the areas of motivation that influence people's continued involvement in learning. Overall, satisfaction refers to the positive feelings a learner associates with the learning experience itself.

The CIS contains 34 items divided into four subscales corresponding to the ARCS model areas of Attention, Relevance, Confidence, and Satisfaction (see Appendix A). The Attention scale consists of eight items, two of which are reverse loaded. The Relevance scale contains nine items, two of which are reverse loaded. The Confidence scale contains eight items with three reverse loaded. The Satisfaction scale comprises nine items with two reverse loaded. Each item has a five-point Likert-scale for responses: 1-not true, 2-slightly true, 3-moderately true, 4-mostly true, and 5-very true (Keller, 2010). While the CIS comprises four subscales, they are well correlated and all designed to measure the construct of situational motivation (Keller, 2010)

With permission from Dr. Keller (Keller, 2011) the CIS was modified to fit the circumstances of the program evaluation. Modifications included changing delivery from paper to electronic and specifying course identifiers (see Appendix B).

Construct validity and reliability. Keller designed the CIS to measure situational motivation. Thus, the CIS is not intended to measure an overall construct or trait of motivation across areas. Rather the CIS is designed to gather specific information regarding a student's motivation in a particular class at a moment in time (Keller, 2010). Keller developed a pool of potential items for CIS based on a review of motivational concepts, measurement instruments, and motivational strategies. A group of ten graduate students who had a strong knowledge of the literature related to motivation reviewed, refined, and reduced the pool items. A second group of ten college and graduate students who were not aware of motivation research responded to the pool items twice. Initially these testers responded as if rating a highly motivating course, then they responded as if they were in a pedantic course. This process identified items that were answered similarly, regardless of whether a course was perceived to be motivating or not. These items were either deleted or modified and re-tested (Keller, 2010). The results of multiple administrations of the CIS led to minor refinements.

Students' CIS scores were compared to course grades, and overall Grade Point Averages (GPAs) to establish criterion validity for the CIS within a narrowly defined situation or context, i.e. within math class. Significant correlations ($p < .05$) between scores on CIS and course grades without significant correlation with GPA ($p > .05$) supported the validity of the CIS as a situation-specific measure of motivation.

Keller's analysis of the internal consistency of the CIS resulted in a Cronbach's alpha of .95. His results for each of the scales resulted in a Cronbach's alpha coefficients for the Attention scale of .84, for the Relevance scale of .84, for the Confidence scale of .81, for the Satisfaction scale of .88 (Keller, 2010). Other studies have used CIS with overall reported Cronbach alpha coefficients ranging from .90-.95 (Carson, 2006; Keller & Suzuki, 2004; Small, Zakaria & El-Figuigul, 2004). These statistics are well within the accepted range for an instrument of this type and indicate that subscales are measuring the same fundamental construct (Nunnaly, 1978). In this study, the Cronbach's alpha for the CIS on pretest was .84 and on posttest was .83. Both pre- and post-test CIS alpha scores were consistent with previous findings and indicated internal consistency and reliability of the CIS.

Attitudes toward mathematics. Students' attitudes toward mathematics were measured using five of the nine sub-scales of the Fennema-Sherman Mathematics Attitudes Scales (1976). The full battery Fennema-Sherman Mathematics Attitudes Scales (FSMAS) consists of 108 items divided into nine scales: (a) Confidence in Learning Mathematics, (b) Attitude toward Success in Mathematics scale, (c) Usefulness of Mathematics scale, (d) Mathematics Anxiety scale, (e) Effectance Motivation scale, (f) The Teacher scale, (g) Mathematics and Sex scale, (h)

Mathematics as a Male Domain scale, and (I) the Mother/Father scale. Each scale comprises twelve statements, six loaded positively, and six loaded negatively. Each statement has a five-point Likert-scale for responses: 1-strongly agree, 2-agree, 3-undecided, 4-disagree, and 5-strongly disagree. The Anxiety scale is reverse scored so that higher scores indicate less anxiety and lower scores indicate more anxiety.

The program evaluation used a modified FSMAS that consisted of 60 items comprising five of the nine scales within the full battery: Confidence in Learning Mathematics scale, Attitude toward Success in Mathematics scale, Usefulness of Mathematics scale, Mathematics Anxiety scale, and Effectance Motivation scale. The program evaluation limited data collection to only those scales that directly addressed the research questions, thus the Teacher scale, Mathematics and Sex scale, Mathematics as a Male Domain scale, and the Mother/Father scale were not used.

Construct validity and reliability. Fennema and Sherman established content validity for the FSMAS by independently developing items for the nine domains and assessing the validity of the items. Fennema and Sherman field-tested the initial 173 items with 367 mathematics students to determine the items with the highest internal correlations. The instrument was then reduced to twelve items for each of the nine domains. Fennema and Sherman administered this revised original version of FSMAS in four Wisconsin high schools and calculated split-half reliabilities for each scale resulting in coefficients ranging from 0.86 to 0.93 (1976). These statistics are well within the accepted range for an instrument of this type and indicate that subscales are measuring the same fundamental construct (Nunnaly, 1978). Over the past thirty-five years, numerous researchers have used FSMAS with reported measures of reliability consistent with Fennema and Sherman's original work (Broadbooks, Elmore, Pedersen, and Bleyer, 1981; Kim & Keller, 2010; Mulhern & Rae, 1998).

Broadbooks, Elmore, Pedersen, and Bleyer (1981) studied the construct validity of the FSMAS. They administered the full battery FSMAS to 1,591 middle school students and the resulting factor analyses indicated that the instrument measured different constructs within the domain of mathematical attitudes. Broadbooks and colleagues' findings suggested that some items, in particular FSMAS scales, were more associated with the factor in a different scale. Additionally, these findings indicated that the items on the Teacher scale of the FSMAS might measure two distinct factors. Despite some differences from Fennema and Sherman's (1976) original results, Broadbooks and colleagues found "evidence to support the theoretical structure of the Fennema-Sherman Mathematics Attitudes Scales" (p.556). They also concluded that the FSMAS was an appropriate scale to measure attitudes toward mathematics for middle school students (1981). Mulhern and Rae (1998) conducted a validity analysis of the nine scales on 196 secondary students in the Republic of Ireland. In addition to confirming validity, they reported Cronbach alpha coefficients for each scale ranging from .83 to .91 with an overall alpha coefficient of .96. For this study, the FSMAS Cronbach's alpha pretest was .82 and posttest was .83. These Cronbach alpha scores were consistent with other studies and indicated inter-item consistency and reliability for the FSMAS.

Data Collection Procedures

The International School of Kenya (ISK) conducted a program evaluation of the instructional effectiveness of using digital games in their middle school mathematics courses. ISK provided ISA student data, students' scores on internal assessments, and survey results.

As part of the program evaluation, ISK Middle school students completed an online pre-treatment survey that contained the CIS and the FSMAS items. Students completed the survey during their regular forty-minute advisory block. ISK middle school advisory groups generally meet four times per week with a faculty member. Faculty members who oversaw the administration of the online survey attended a short informational session to review the directions for administering the survey. After the nine-week treatment, students completed an online post-treatment survey containing the CIS and FSMAS items during an advisory block. In addition, on the post-treatment survey, students were asked to indicate the amount of time per week on average they spent playing *Labyrinth* outside of class. ISK math teachers provided internal math exam scores for students pre and post treatment. The results and analysis of pre and post treatment survey and achievement data are reported in chapter four.

Data Analysis

This study examined data collected during a program evaluation. The data of interest concerned the effect of digital games and sex on three dependent variables (a) achievement, (b) motivation, and (c) attitude. A Pearson Correlation of 30 randomly selected students' ISK pre-treatment exam scores and their October 2011 ISA scores showed a strong a strong relationship (r = .82), which established criterion validity for ISK exams as a measure of mathematical achievement. Cronbach's alphas for the CIS (pretest α = .84, posttest α = .83) and FSMAS (pretest α = .82, posttest α = .83) were consistent with previous findings and confirm reliability of both instruments.

Random assignment to group should ensure that the groups did not differ significantly. However, the pre-treatment means on each variable for the control and treatment groups were examined (see Table 1) and a *t*-test comparing the means for the two groups was performed which confirmed their baseline scores did not differ significantly (see Table 2).

Table 1

Mean, Standard Deviation, and Standard Error of the Mean for Control and Treatment Pretest Scores

Pretest Scale	Group	*N*	*M*	*SD*	*SEM*
Achievement	Control	84	80.99	10.755	1.173
	Treatment	84	81.11	11.237	1.226
CIS Attention	Control	84	22.15	2.695	.294
	Treatment	84	22.07	2.723	.297
CIS Relevance	Control	84	25.93	4.219	.460
	Treatment	84	25.89	4.294	.468
CIS Confidence	Control	84	28.18	4.867	.531
	Treatment	84	27.94	4.624	.505
CIS Satisfaction	Control	84	27.86	3.275	.357
	Treatment	84	27.64	3.400	.371
FS Confidence	Control	84	37.23	8.014	.874
	Treatment	84	37.58	8.244	.899
FS Success	Control	84	52.36	3.266	.356
	Treatment	84	52.10	2.960	.323
FS Usefulness	Control	84	27.11	9.527	1.040
	Treatment	84	27.04	9.458	1.032
FS Anxiety	Control	84	34.21	3.716	.405
	Treatment	84	33.70	3.914	.427
FS Effectance	Control	84	30.86	9.151	.998
	Treatment	84	31.25	9.132	.996

Table 2

Independent Samples Test: t-test for Equality of Means of Pretest Scores between Control and Treatment Groups

Pretest Scales	*t*	*Df*	*Sig. (2-tailed)*	*Mean Difference*
Achievement	-.070	166	.944	-.119
CIS Attention	.199	166	.842	.083
CIS Relevance	.054	166	.957	.036
CIS Confidence	.325	166	.746	.238
CIS Satisfaction	.416	166	.678	.214
FS Confidence	-.285	166	.776	-.357
FS Attitude Toward Success	.545	166	.587	.262
FS Usefulness	.049	166	.961	.071
FS Anxiety	.869	166	.386	.512
FS Effectance	-.279	166	.781	-.393

Note. CIS = Course Interest Survey, FS = Fennema-Sherman Mathematics Attitudes Scales

The plan for multivariate analysis of covariance (MANCOVA) resulted in six Models corresponding to each of the six research questions. Table 3 shows the predictor variables, covariates, and outcome variables for each Model. For each Model, the data were examined to ensure it met the series of assumptions that underlie MANCOVA. The results of the tests of assumptions are reported in Appendix D. Models 1 through 2B were univariate models. Levene's tests confirmed that Models 1 through 2B met the assumption of equal variances. Models 1through 2B also had equal pretest means and showed normally distributed residuals, as indicated by a normal distribution on the Q-Q plot. Therefore, all assumptions were met for these models.

Table 3

Models, Associated Research Questions, Predictor Variables, Dependent Variables, Covariates, and Reference Group

Model	Research Question	Predictor Variable(s)	Dependent Variable(s)	Covariate(s)	Reference Group
1	What effects did playing *Lure of the Labyrinth* have on middle school students' math achievement?	Treatment	Post achievement	Pre Achievement	Control
2A	Did the change in achievement scores differ for boy versus girls?	Treatment Sex	Post achievement	Pre Achievement	Control Male
2B	Were the effects of playing *Labyrinth* on middle school students' math achievement different for boys versus girls?	Treatment Sex Treatment*Sex	Post achievement	Pre Achievement	Control Male
3	What effects did playing *Labyrinth* have on middle school students' motivation to learn math?	Treatment	Post Attention (CIS) Post Relevance (CIS) Post Confidence (CIS) Post Satisfaction (CIS)	Pre Attention (CIS) Pre Relevance (CIS) Pre Confidence (CIS) Pre Satisfaction (CIS)	Control
4A	Did the change in motivation scores differ for boy versus girls?	Treatment Sex	Post Attention (CIS) Post Relevance (CIS) Post Confidence (CIS) Post Satisfaction (CIS)	Pre Attention (CIS) Pre Relevance (CIS) Pre Confidence (CIS) Pre Satisfaction (CIS)	Control Male
4B	Did the effects of playing *Labyrinth* on middle school students' motivation to	Treatment Sex	Post Attention (CIS) Post Relevance (CIS)	Pre Attention (CIS) Pre Relevance (CIS)	Control Male

	learn math differ for boys versus girls?	Treatment*Sex	Post Confidence (CIS) Post Satisfaction (CIS)	Pre Confidence (CIS) Pre Satisfaction (CIS)	
5	Did playing *Labyrinth* affect middle school students' attitudes toward math?	Treatment	Post Confidence (FS) Post Success (FS) Post Usefulness (FS) Post Anxiety (FS) Post Effectance (FS)	Pre Confidence (FS) Pre Success (FS) Pre Usefulness (FS) Pre Anxiety (FS) Pre Effectance (FS)	Control
6A	Did the change in attitude scores differ for boy versus girls?	Treatment Sex	Post Confidence (FS) Post Success (FS) Post Usefulness (FS) Post Anxiety (FS) Post Effectance (FS)	Pre Confidence (FS) Pre Success (FS) Pre Usefulness (FS) Pre Anxiety (FS) Pre Effectance (FS)	Control Male
6B	Did the effects of *Labyrinth* on middle school students' attitudes toward math differ for boys versus girls?	Treatment Sex Treatment*Sex	Post Confidence (FS) Post Success (FS) Post Usefulness (FS) Post Anxiety (FS) Post Effectance (FS)	Pre Confidence (FS) Pre Success (FS) Pre Usefulness (FS) Pre Anxiety (FS) Pre Effectance (FS)	Control Male

Models 3 through 6B were multivariate models. In all but three cases, Levene's test indicated the assumption of equal variances was met. However, because the sample size was relatively large, the test was robust to violations of this assumption (Norusis, 1999). In all cases for Models 3 through 6B, Box's M was non-significant which indicated the covariance matrices were equal and this underlying MANCOVA assumption was met. Bartlett's test of sphericity for Models 3 through 6B, indicated the dependent variables were significantly correlated among themselves and thus the assumption was met. In all cases for Models 3 through 6B, the mean pretest scores were equal across groups, which met the assumption. In 18 out of 31 cases in Models 3 through 6B, Q-Q plots showed the residuals were normally distributed, indicating that the assumption was met. However, 42% of the tests revealed a non-normal Q-Q plot. Although the assumptions was not met in these cases, the relatively large sample sizes rendered the tests robust to violations of the assumption (Norusis, 1999).

Conclusion validity. Like all statistical measures, MANCOVA are subject to type I and II errors. Type I errors, in which the null hypothesis is incorrectly rejected, were minimized by the experimental research design of the program evaluation and the establishment of an alpha level of .05 to indicate significance. Thus, the likelihood that the null hypothesis was rejected in error is less than 5%. Type II errors, in which the null hypothesis is accepted in error, were minimized in by the relatively large sample size (N = 168) of the data set and computing effect size based on alpha .05 and power of .8. Treatment fidelity was established by analyzing student responses to the post treatment survey question of whether they had played *Labyrinth* outside of class and if so, to estimate time engaged with *Labyrinth* outside of class. Only 3 of the 84 students in the control group indicated they had accessed *Labyrinth* outside of class (none had accessed it within class). The mean playtime of these three students was 8.3 minutes. The low number of students in the control group who played *Labyrinth* (n = 3) and the small amount of exposure to the game (M= 8.3 minutes) was not interpreted as a threat to the overall research design.

Summary

This study analyzed data collected during a program evaluation. The analysis explored the effects of playing a digital game on 168 middle school students' mathematical achievement, attitudes toward mathematics, and motivation in mathematics class based on data collected during a school program evaluation. Students in the program evaluation were assigned to treatment and control groups based on random stratified sampling to ensure equal representation of sex and grade level. Mathematical achievement was measured using internal standards-based exams. Mathematical attitude was measured using five of the nine Fennema-Sherman Mathematics Attitudes Scales (FSMAS). Motivation was measured using the Course Inventory System (CIS). Multivariate analyses of covariance (MANCOVA) was used to correlate change in mathematical achievement, attitudes and motivation with the predictor variables of group (control and treatment), sex (male and female), and the interaction of treatment and sex.

Chapter 4

Findings

Multivariate analysis of covariance (MANCOVA) indicated that for both boys and girls playing *Lure of the Labyrinth* had a significant, positive, small effect on mathematical achievement; a significant, positive, moderate effect on motivation; and a significant, positive, small effect on attitudes. However, the positive effect of playing *Lure* was not found in all measures of motivation and attitude. Further, analysis of the subscale measures of motivation and attitude indicated some differences between the effects for boys and girls. The following sections present the data analysis for each research question.

Research Question 1: What effects did playing *Lure of the Labyrinth* have on middle school students' mathematical achievement?

In order to answer each research question, the researcher used multivariate analysis of covariance (MANCOVA). The MANCOVA related to the first research question entailed regressing the posttest math achievement scores on the covariate of pretest math achievement scores and on treatment as a predictor variable. Establishing pretest scores as the covariate in essence removed the variation in math achievement scores and enabled the researcher to compare the change in math achievement scores from pre- to post-test for the control versus the treatment group.

As shown in Table 4, the 84 control group students exhibited an average math score decline of -.38 point from 80.99 points (*SD* = 10.76) at pretest to 80.61 points (*SD* = 11.10) at posttest. In contrast, the 84 students in the treatment group improved their mathematical achievement scores + 1.49 points from an average of 81.11 points at pretest (*SD* = 11.24) to 82.60 points at posttest (*SD* = 11.15).

Table 4

Achievement Score Means at Pre- and Post-test for Students in the Treatment and Control Groups

	Achievement Pretest		Achievement Posttest	
Group	Mean	*SD*	Mean	*SD*
Control (*N* =84)	80.99	10.76	80.61	11.10
Treatment (*N* =84)	81.11	11.24	82.60	11.15
Total (*N* =168)	81.05	10.97	81.60	11.14

In Table 5, Model 1 the second *b* coefficient is -1.87, which is associated with the predictor variable labeled as treatment. The regression coefficient, also referred to as the *b* coefficient, was derived by calculating the difference between the change of scores from pre- to post-test for the control group (-.38) and the change in scores from pre- to post-test for the treatment group (+1.49). Using the control group as the reference, the difference in the change of the scores is -1.87 (-.38 - 1.49). This regression coefficient ($b = -1.87$) is significant ($p < .0005$) which indicates that the differential growth on mathematical achievement for treatment group students and control group students is likely due to the effects of the treatment, playing *Labyrinth*.

The mean gain for the treatment group from pre- to post-test was 1.49 points (from 81.11 at pretest to 82.60 at posttest). The pooled standard deviation was 11.19 (11.24 + 11.15 = 22.39/2 = 11.19, with rounding error). Cohen's *d* was used to calculate effect size (ES) based on the differences in the means (1.49) divided by the pooled standard deviation (11.19) resulting in an ES of .13. Although statisticians differ on the guidelines for interpreting effect size, Cohen (1988) suggested viewing an effect size less than .3 as small, between .3 and .7 as moderate, and an effect size larger than .7 as large . Thus, it seems that playing *Labyrinth* had a significant ($p <$ .0005) but small effect (ES = .13) on students' mathematical achievement.

Table 5

Regression Coefficients for Predictors of Math Achievement Scores

Predictors	Model 1	Model 2A	Model 2B
Achieve pre	.97***	.97***	.97***
Treatment	-1.87***	-1.87***	1.10
Sex		.50	1.27+
Treatment*Sex			-1.54

Note. Achieve = Mathematical Achievement; Pre = pretest
+p < .10. *p < .05. **p < .005. ***p < .0005.

Research Question 2: Were the effects of playing *Labyrinth* on middle school students' mathematical achievement different for boys versus girls?

Table 6 shows the means and standard deviations for pre and post math achievement scores for boys and girls in the control and treatment groups. The difference in the change scores from pre- to post-test for boys versus girls was analyzed by MANCOVA. Table 5, Model 2A shows the regression coefficient for posttest mathematical achievement with the predictor variable of sex when pretest achievement scores were the covariate. The regression coefficient (b =.50) was not significant (p = *ns*) which indicated boys and girls showed no meaningful differences in their growth in math achievement from pre- to post-test. As shown in Table 5, Model 2B, the regression coefficient for the interaction term between treatment and sex also indicated no significant differential effect of treatment for boys and girls (b = -1.54, p = *ns*). Thus, it seems playing *Labyrinth* had a positive but small effect on math achievement for both girls and boys (b = -1.87, $p < .0005$, ES = .13) with no significant difference of treatment due to sex (b = -1.54, p = *ns*).

Table 6

Means and Standard Deviations on Pre and Post Achievement Scores

	Achievement pre		Achievement post	
Sex-Group (*N*)	Mean	SD	Mean	SD
Boys-Control (*N* = 42)	80.45	10.465	79.95	10.815
Boys-Treatment (*N* = 42)	80.43	11.190	82.57	11.297
Boy-Total (*N* = 84)	80.44	10.768	81.26	11.071
Girl-Control (*N* = 42)	81.52	11.138	81.26	11.477
Girl-Treatment (*N* = 42)	81.79	11.377	82.62	11.129
Girl-Total (*N* = 84)	81.65	11.191	81.94	11.256
Control (*N* = 84)	80.99	10.755	80.61	11.103
Treatment (*N* = 84)	81.11	11.237	82.60	11.145
Total (*N* = 168)	81.05	10.966	81.60	11.136

Research Question 3: What effects did playing *Labyrinth* have on middle school students' motivation to learn mathematics?

The study employed four measures of situational motivation from the CIS: attention, relevance, confidence, and satisfaction. Table 7 shows the means and standard deviations on pre- and post-test scores for the attention measure.

Table 7

Means and Standard Deviations on Pre- and Post-Test Attention Scores (CIS)

	Attention pre		Attention post	
Sex-Group (*N*)	Mean	SD	Mean	SD
Boys-Control (*N* = 42)	21.90	2.487	22.07	2.645
Boys-Treatment (*N* = 42)	21.81	2.725	23.48	2.830
Boy-Total (*N* = 84)	21.86	2.594	22.77	2.813
Girl-Control (*N* = 42)	22.40	2.897	22.17	3.131
Girl-Treatment (*N* = 42)	22.33	2.729	22.95	2.793
Girl-Total (*N* = 84)	22.37	2.798	22.12	2.975
Control (*N* = 84)	22.15	2.695	22.12	2.881
Treatment (*N* = 84)	22.07	2.723	23.21	2.807
Total (*N* = 168)	22.11	2.702	22.67	2.888

Overall, students in the control group lost .03 points on their attention scores from pre- to post-test whereas the treatment group gained 1.14 points. Using the control group's change in attention score as the reference, the overall change between the two groups is -1.17 (-0.3 - 1.14 = -1.17). Thus, the students in the control group at posttest gained an average of 1.17 fewer points on their self-reporting of Attention than the students in the treatment group. Table 8, Model 3 shows the regression coefficient for posttest attention (b = -1.17) and indicates the change in attention scores is significantly different for the control versus treatment groups (p < .0005). Since pretest attention scores were used as a covariate, it is probable that the significant difference in posttest attention scores is attributable to the effects of treatment, playing *Labyrinth*. Interpreting the effect (1.17 points) in the context of the pooled standard deviation of

Attention (2.8 points, with rounding error) yields an effect size of .42. According to Cohen's guidelines as explained previously, treatment had a moderate sized effect on growth in Attention.

Regression coefficients for the three other CIS measures of motivation (relevance, confidence, and satisfaction) are also shown in Table 8, Model 3. Analysis of change scores from pre- to post-test for the three other measures of situational motivation when pretest scores were used as covariates found no significant effects of treatment. Means and standard deviations for scores on the relevance, confidence, and satisfaction measures of the CIS are shown in Appendix E.

Table 8

Regression Coefficients for Predictors of Math Motivation Scores

Predictors	Model 3	Model 4A	Model 4B
	Attention	Attention	Attention
	Post	Post	Post
Attention (pre)	.92***	.95***	.95***
Relevance (pre)	.05	.07*	.06+
Confidence (pre)	-.05	-.07+	-.06+
Satisfaction (pre)	.03	.02	.02
Treatment	-1.17***	-1.17***	-.91**
Sex		.78***	1.03***
Treatment*Sex			-.51
	Relevance	Relevance	Relevance
	Post	Post	Post
Attention (pre)	.03	.02	.02
Relevance (pre)	.90***	.89***	.89***

Confidence (pre)	.02	.03	.03
Satisfaction (pre)	.05	.05	.06
Treatment	-.22	-.22	-.24
Sex		-.13	-.15
Treatment*Sex			.05
	Confidence Post	Confidence Post	Confidence Post
Attention (pre)	.02	.02	.02
Relevance (pre)	.11*	.11*	.12**
Confidence (pre)	.91***	.91***	.91***
Satisfaction (pre)	-.06	-.06	-.06
Treatment	-.09	-.09	-.38
Sex		-.04	-.32
Treatment*Sex			.58
	Satisfaction Post	Satisfaction Post	Satisfaction Post
Attention (pre)	.01	.01	.01
Relevance (pre)	.08	.09+	.10*
Confidence (pre)	.05	.04	.03
Satisfaction (pre)	.74***	.73***	.74***
Treatment	.05	.06	-.52
Sex		.20	-.36
Treatment*Sex			1.15*

Note. Pre = pretest; Post = posttest

+p < .10. *p < .05. **p < .005. ***p < .0005.

Research Question 4: Did the effects of playing *Labyrinth* on middle school students' motivation to learn mathematics differ for boys versus girls?

The study used the CIS's four measures of situational motivation: attention, relevance, confidence, and satisfaction. Table 7 shows the means and standard deviations on pre- and post-test scores for the attention measure. The difference in mean scores at posttest compared to pretest for boys (22.77 – 21.86) showed a posttest gain of .91 points on the attention measure. Whereas, the difference in mean scores at posttest compared to pretest for girls (22.56 - 22.37) showed a posttest gain of .19 points on the attention measure. Thus, boys gained an average of .72 points more than girls did (.91 -. 19) on the posttest attention measure. Model 4A on Table 8, shows the regression coefficients for each of the measures of situational motivation with the predictor variables. Within the attention measure of situational motivation, the regression coefficient (b =.78) indicated that the growth of attention scores is significantly different for boys versus girls ($p < .0005$). Thus, sex (boy or girl) significantly affected the amount of change from pre- to post-test on the attention measure. The size of the effect of sex on attitude scores was determined by taking the difference between boys' gains and girls' gains (.91 -. 19 = .72) and dividing it by the pooled standard deviation of pre- and post-test scores (2.79), yielding an effect size of .26 (.72 divided by 2.79). Thus, it seems there was a small (Cohen's d = .26) but significant (b =.78, $p < .0005$) effect of sex on attention scores. In Table 8, Model 4A also shows that the regression coefficients for the three other measures of situational motivation (relevance, confidence and satisfaction) did not indicate a differential effect of sex when pretest scores were used as covariates. Means and standard deviations for scores on the relevance, confidence, and satisfaction measures of the CIS are shown in Appendix E.

Models labeled 4B examined the joint effect of Treatment and Sex. Although treatment had a moderate effect on attention scores (b = -1.17, $p < .0005$, ES = .42) and sex had a small effect on attention scores (b =.78, $p < .0005$, ES = .26), treatment did not have a differential impact on attention for boys and girls (b = -.51, p = *ns*). Figure 5 shows the pre- and post-test means for boys in the control group (blue line) and indicates a slight gain; whereas girls in the control group (red line) show a slight decline from pre- to post-test. Figure 5 also shows the means for pre- and post-test for boys in the treatment group (green line) and girls (purple lines) and reflects the moderate sized gains in attention related to treatment. In sum, the students in the treatment group showed moderately greater growth on attention scores than students in the control group, and boys gained slightly more than girls did. However, the differential effects for boys and girls in the treatment group versus boys and girls in the control group were not significant.

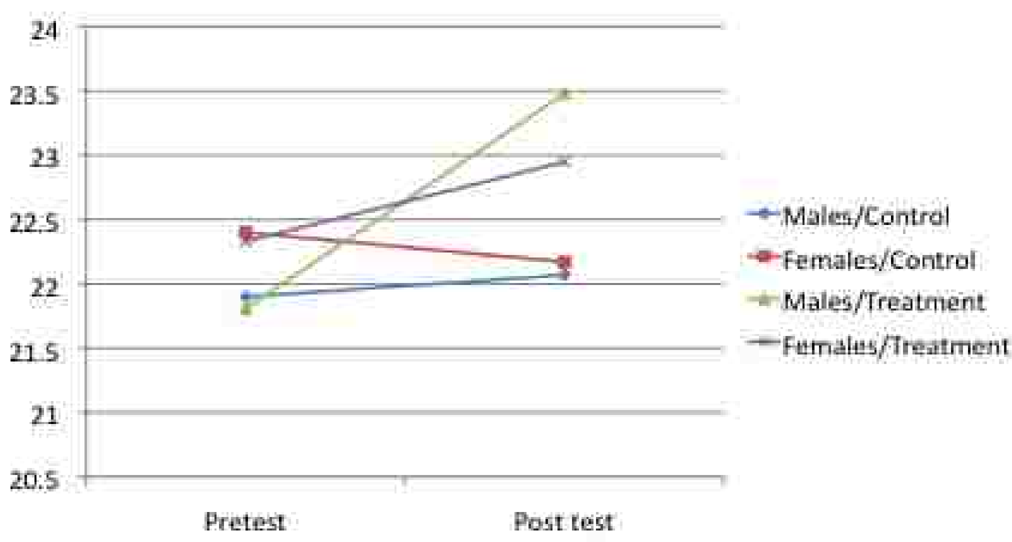

Figure 5. Attention scores pre- and post-test for boys and girls in the treatment and control groups. This figure shows an increase in Attention scores for boys in the treatment (green line) boys in the control groups (blue line) and girls in the treatment group (purple line). However, a decrease in Attention scores occurred for the girls in the control group (red line).

Model 4B on Table 8 also shows that for Relevance and Confidence, students overall evidenced significant gains from pre- to post-test. However, these effects did not appear to be linked to sex or to the joint effect of treatment and sex In other words, boys and girls in both treatment and control groups progressed at the same pace with relation to Relevance and Confidence regardless of whether they played *Labyrinth* or not. Means and standard deviations for scores on the relevance and confidence measures of the CIS are shown in Appendix E.

Although playing *Labyrinth*, the treatment in this study, did not have a significant effect on the satisfaction measure of motivation (Table 8, Model 3, $b = .05$, $p = ns$), within the treatment group the pattern of change in the satisfaction measure for boys was different from the pattern of change for girls. Model 4B on Table 8 shows the regression coefficient for the joint interaction of treatment and sex on satisfaction ($b = 1.15$), which is significant ($p < .05$) and shows a small (ES = .09) positive effect for girls in the treatment group. Table 9 shows the means for each group at pre- and post-test on satisfaction. Figure 6 is a line graph showing the pattern of change in mean satisfaction scores from pre- to post-test for each group. Figure 6 shows that the boys in the control group (blue line) gained on their satisfaction scores whereas the girls in the control group (red line) demonstrated a decline in satisfaction scores. Figure 6 also shows the decline in satisfaction scores from pre- to post-test for boys in the treatment group (green line) and the gain in satisfaction for girls in the treatment group (purple line).

Table 9

Pretest and Posttest Means on CIS Satisfaction by Group and Sex

Group (Sex)	Pretest	Posttest
Control (Boy)	27.79	28.17
Control (Girl)	27.93	27.67
Treatment (Boy)	27.83	27.64
Treatment (Girl)	27.45	27.74

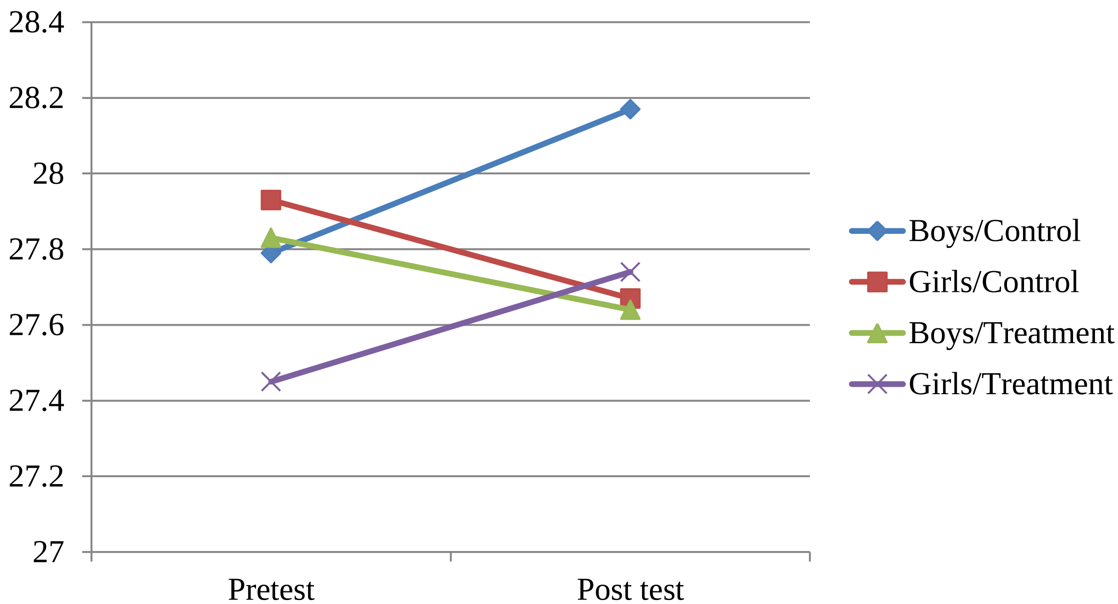

Figure 6. Change on CIS Satisfaction mean score from pretest to posttest for boys and girls in the treatment group and control group.

Research Question 5: Did playing *Labyrinth* affect middle school students' attitudes toward mathematics?

The modified version of the FSMAS used in this study provided five measurements of students' attitudes toward mathematics: Confidence, Attitude Toward Success, Usefulness, Anxiety, and Effectance. Each measurement was addressed in a separate part of Model 5 in Table 10. The regression coefficient for the posttest confidence scores shows that when pretest confidence scores are used as a covariate, treatment is not a predictor of change in confidence in mathematics scores (b = -.23, p = *ns*). Similarly, the regression coefficient for the satisfaction score indicates that treatment, when pretest satisfaction scores are controlled for, was not a significant predicator of the change in attitude toward success scores (b = .12, p = *ns*). Thus, treatment, playing *Labyrinth*, did not have a differential effect on either confidence or attitude toward success measurements of mathematical attitudes. Means and standard deviations for the FSMAS confidence and success scores are shown in Appendix F.

Table 10

Regression Coefficients for Predictors of Math Attitude Scores

Predictors	Model 5 Confidence Post	Model 6A Confidence Post	Model 6B Confidence Post
Confidence (pre)	.90***	.90***	.90***
Success (pre)	-.01	-.01	-.01
Usefulness (pre)	.06*	.06*	.06*
Anxiety (pre)	-.05	-.05	-.05
Effectance (pre)	.04+	.04+	.04+
Treatment	-.23	-.23	-.29
Sex		-.28	-.33
Treatment*Sex			.11
	Success Post	Success Post	Success Post

Confidence (pre)	.05*	.05*	.05*
Success (pre)	.82***	.82***	.82***
Usefulness (pre)	-.01	-.01	-.01
Anxiety (pre)	.02	.02	.02
Effectance (pre)	-.02	-.02	-.02
Treatment	.12	.12	-.05
Sex		-.26	-.43
Treatment*Sex			.34
	Usefulness Post	Usefulness Post	Usefulness Post
Confidence (pre)	.09*	.09*	.09*
Success (pre)	-.04	-.04	-.04
Usefulness (pre)	.87***	.87***	.87***
Anxiety (pre)	.04	.04	.04
Effectance (pre)	.04	.04	.04
Treatment	-.77*	-.77*	-.87
Sex		.32	.22
Treatment*Sex			.20
	Anxiety Post	Anxiety Post	Anxiety Post
Confidence (pre)	.03	.03	.03
Success (pre)	-.03	-.03	-.03
Usefulness (pre)	.01	.01	.01
Anxiety (pre)	.86***	.86***	.86***
Effectance (pre)	.01	.01	.01
Treatment	-.11	-.11	-.66*
Sex		-.15	-.69*
Treatment*Sex			1.10*
	Effectance Post	Effectance Post	Effectance Post
Confidence (pre)	.04	.03	.04
Success (pre)	.02	.02	.02
Usefulness (pre)	.03	.03	.03
Anxiety (pre)	.04	.04	.04
Effectance (pre)	.90***	.90***	.90***
Treatment	-1.18***	-1.18***	-.96**
Sex		.33	.55
Treatment*Sex			-.44

Note. Pre = pretest; Post = posttest

+p < .10. *p < .05. **p < .005. ***p < .0005.

The MANCOVA results for Model 5 on Table 10 showed that for Usefulness of Math the effect of treatment was significant when the usefulness pretest scores were used as a covariate ($b = -.77$, $p < .05$). Table 11 shows the pre- and post-test mean FSMAS usefulness of mathematics scores and their standard deviations. The treatment group showed a gain of .23 points on the Usefulness of Mathematics score from pretest ($M = 27.04$) to posttest ($M = 27.27$). Alternately, the control group showed a decline of .59 points in the FSMAS Usefulness of Math scores from pretest ($M = 27.11$) to posttest ($M = 26.52$). Thus, the difference between the treatment group gain and the control group loss was .82 points (.23- [-.59] = .82). As with previous significant effects, Cohen's *d* was used to calculate the effect size by dividing the difference in the scores between the treatment and control groups (.82) by the pooled standard deviation (9.4, with rounding error). Using Cohen's (1988) interpretation of the size of effects, the resulting Cohen's *d* value of .09 would be considered a small effect of treatment on the FSMAS Usefulness of Mathematics scores.

Table 11

Means and Standard Deviations on Usefulness of Mathematics for Treatment Group and Control Group at Pretest and Posttest

	Usefulness Pretest		Usefulness Posttest	
Group	Mean	*SD*	Mean	*SD*
Control ($n = 84$)	27.11	9.53	26.52	9.47
Treatment ($n = 84$)	27.04	9.46	27.27	9.23
Total ($n = 168$)	27.07	9.46	26.90	9.33

The regression coefficient for treatment as a predictor of FSMAS Mathematical Anxiety posttest scores ($b = -.11$, $p = ns$), indicated that when pretest anxiety scores were controlled for, playing *Labyrinth* did not have a significant effect on overall posttest Anxiety scores. Means and standard deviations for Anxiety scores are reported in Table 13.

Model 5 on Table 10 shows that for posttest score of the FSMAS Effectance measure of mathematical attitudes, treatment was a significant predictor (b = -1.18, p < .0005) when pretest scores were used as a covariate. To evaluate the size of the significant effect of playing Labyrinth on the FSMAS Effectance measure of math attitudes, the means and standard deviations of the scores were examined. Table 12 shows that the mean FSMAS Effectance score for the control group decreased .23 points from pretest (M = 30.86) to posttest (M = 30.63). In contrast, students in the treatment group gained .90 points on the mean FSMAS Effectance score from pretest (M = 31.25) to posttest (M = 32.15). Thus, the overall difference in the change of Effectance scores between the treatment and control groups was 1.13 points (.90 – [-.23] = 1.13). Evaluating the difference in the treatment and control groups' effectance scores in the context of the pooled standard deviation, which was 8.9 with rounding error, yielded an effect size of about .13. Therefore, playing Labyrinth has a small (Cohen's d = .13) but significant (b = -1.18, p < .0005) effect on the FSMAS Effectance measure of mathematical attitudes.

Table 12

Means and Standard Deviations on Effectance for Treatment Group and Control Group at Pretest and Posttest

	Effectance Pretest		Effectance Posttest	
Group	Mean	*SD*	Mean	*SD*
Control (n = 84)	30.86	9.15	30.63	8.91
Treatment (n = 84)	31.25	9.132	32.15	8.764
Total (n = 168)	31.05	9.116	31.39	8.843

Research Question 6: Did the effects of *Labyrinth* on middle school students' attitudes toward mathematics differ for boys versus girls?

Model 6A in Table 10, shows that when pretest scores were used as covariates, sex is not a predictor of the posttest score for any of the five FSMAS measures of attitudes toward mathematics. Therefore, boys and girls showed similar patterns of change from pre- to post-test on the FSMAS measures of confidence, success, usefulness, anxiety, and effectance.

Models 6B in Table 10 used the joint effect of Sex and Treatment as the predictor of outcome scores on the five measures of mathematical attitudes and the pretest scores as the covariates. The interaction effect was not significant for four of the five attitudes: Confidence, Success, Usefulness, and Effectance. However, a significant joint effect of treatment and sex occurred on the FSMAS Anxiety measure of attitudes toward math (b = 1.10, $p < .05$). The effect size was small (ES = .08). Table 13 shows the means and standard deviations on Anxiety scores for the different groups. Boys in the control group showed no change in Anxiety from pre- to post-test. Girls in the control group showed an increase in anxiety as reflected by the decrease on the reverse scored Anxiety scale from pretest (M = 34.21) to posttest (M = 33.79). Boys in the treatment group also showed a decline in Anxiety scores (indicating an increase in math related anxiety) from pretest (M = 33.62) to posttest (M = 33.29). However, girls in the treatment group showed an improvement in anxiety scores from pretest (M = 33.79) to posttest (M = 34.10) indicating a small (ES = .08) decrease in their math related anxiety. Figure 7 is a line graph showing the change on anxiety scores from pre- to post-test. Figure 7 shows no change from pre- to post-test for the control group boys (blue line) the decline in anxiety scores for the girls in the control group (red line) and boys in the treatment group (green line). Figure 7 also shows the improvement in Anxiety scores from pre- to post-test for girls in the treatment group (purple line).

Table 13

Means and Standard Deviations on Anxiety for Boys and Girls in the Control and Treatment Groups at Pre- and Post-Test

		Anxiety Pretest		Anxiety Posttest	
Group	Sex	Mean	*SD*	Mean	*SD*
Control	Boys (*n* = 42)	34.21	3.71	34.21	3.60
	Girls (*n* = 42)	34.21	3.76	33.79	3.69
	Total (*n* = 84)	34.21	3.72	34.00	3.63
Treatment	Boys (*n* = 42)	33.62	3.73	33.29	3.64
	Girls (*n* = 42)	33.79	4.13	34.10	4.04
	Total (*n* = 84)	33.70	3.91	33.69	3.84
Total	Boys (*n* = 84)	33.92	3.71	33.75	3.63
	Girls (*n* = 84)	34.00	3.93	33.94	3.84
	Total (*n* = 168)	33.96	3.81	33.85	3.73

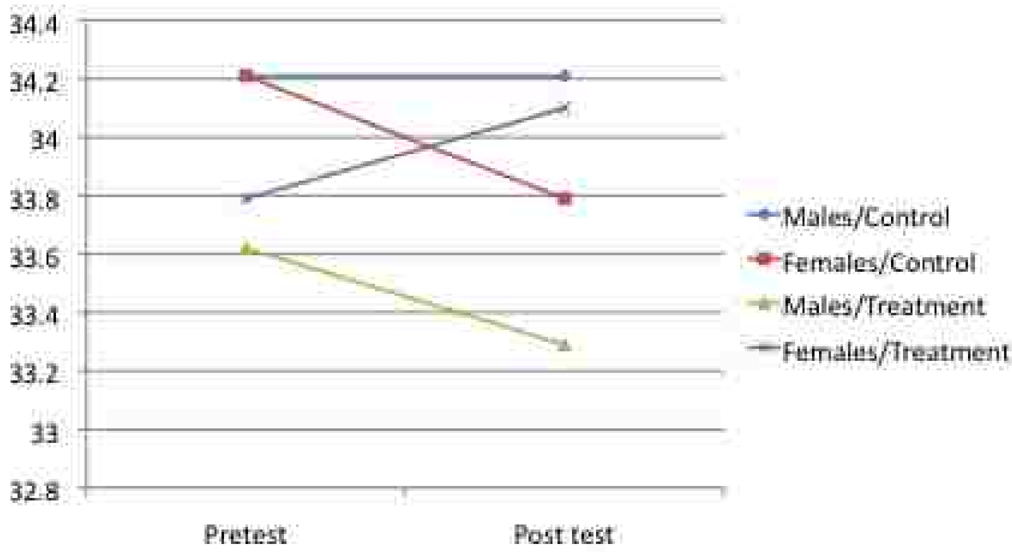

Figure 7. Differential patterns of growth on Anxiety for girls versus boys in the treatment group versus the control group.

Summary

Multivariate analysis of covariance (MANCOVA) of posttest achievement scores when pretest scores were the covariate indicated playing *Lure of the Labyrinth* had a positive, significant (b = -1.87, $p < .0005$) small effect (ES = .13) on both boys' and girls' mathematical achievement. Of the four measures of situational motivation examined by the CIS, MANCOVA results indicated the attention scores showed a significant (b = -1.17, $p < .0005$), moderate effect (ES = .42) of treatment. Further, MANCOVA results indicated that while the growth of attention scores was significantly different for boys versus girls (b =.78, $p < .0005$, ES = .26), treatment did not have a differential impact on attention for boys and girls (b = -.51, p = *ns*). Although playing *Labyrinth*, did not have a significant effect on the satisfaction measure of motivation (b = .05, p = *ns*), within the treatment group boys showed a decline in the satisfaction measure whereas girls showed an increase in satisfaction (b = 1.15, $p < .05$). MANCOVA results indicated a significant attitude change on the Usefulness measure of the FSMAS (b = -.77, p <

.05, ES = .09). In addition, treatment had a significant positive effect on the Effectance measure of math attitudes (*b* = -1.18, *p* < .0005, ES = .13). Changes in scores for boys and girls on all five FSMAS measures of math attitudes were not significantly different. However, MANCOVA results showed a significant joint effect of sex and treatment on the Anxiety measure of attitudes, in which girls who played Labyrinth showed improved scores related to their perception of math related anxiety (*b* = 1.10, *p* < .05, ES = .08) as compared with girls in the control group or boys in the treatment group.

Chapter 5

Discussion

Chapter five discusses the research findings presented in Chapter four. It is divided into seven sections (a) mathematics achievement; (b) mathematics motivation; (c) attitudes toward mathematics; (d) limitations; (e) contributions to research; (f) contributions to practice; and (g) conclusion. The first section discusses treatment effects on participants' mathematics achievement as well as differences in effects for boys versus girls. The second section discusses the effects of playing *Labyrinth* on measures of situational motivation in relationship to math class as well as sex differences. The third section discusses the results of the five subscales of the FSMAS in relation to treatment as well as differences in effects for boys versus girls. The fourth section lists the limitations of this study. The fifth and sixth sections explore this study's contributions to research and practice. The last section summarizes the conclusions presented and provides recommendations for future research.

Mathematics Achievement and *Lure of the Labyrinth*

The first research hypothesis proposed that no significant difference existed between learners' mathematical achievement in the experimental group, who played *Labyrinth*, versus the control group, who did not play. A MANCOVA was performed in which pretest results on benchmarked internal exams were used as the covariate to control for the initial differences among the participants' and students' posttest scores on benchmark exams were the dependent variable. As expected, students who did well at pretest also did well at posttest and students who performed poorly at pretest also struggled at posttest ($b = .97$, $p < .0005$). In addition a significant regression coefficient was associated with playing *Labyrinth* ($b = -1.87$, $p < .0005$).

Since the treatment group was the reference group, the regression coefficient (*b*) means that treatment group students gained an average of 1.87 achievement points more than control group students from pre- to post-test. Since this effect was due to chance only 5 out of 10,000 times, these results indicated that the achievement gains for the treatment group were likely due to the effects of playing *Lure of the Labyrinth.* Effect sizes were judged in the context of the pooled standard deviation of the scores. The calculated effect size for the treatment was 1.49/11.19 = .13, which represents a small effect.

Models 2A and 2B addressed whether the effect on achievement was different for boys versus girls. Model 2A was a stepping stone model and showed again that the treatment effect was significant (b =-1.87). However, the effect of Sex was not significant (b =.50, p = *ns*) which indicated similar patterns of change for girls' growth and boys' growth in math achievement from pre- to post-test. In Model 2B, the interaction term between Treatment and Sex was examined and the non-significant regression coefficient (b = -1.54, p = *ns*) indicated that treatment did not affect boys and girls differently.

The current study contributes further insights to the existing literature on the effectiveness of digital games for learning. This study tested the effects of a serious digital game that included opportunity for collaborative play, as described in Chapter Three, while most previous empirical studies used single player games that had not necessarily been designed based on constructivist pedagogy and may not have included opportunities for collaboration (Ke & Grabowski, 2007; Lee, Luchini, Michael, Norris, & Soloway, 2004; Shaffer, 1997). While students in the treatment group used the TPC, *Labyrinth's* internal game communicator, as well as having freedom to discuss game play in real space, collaboration was not a formal variable of interest in this study. Future research is needed to explore the role of collaboration on math learning outcomes.

Research has indicated differences between boys and girls related to frequency of play of digital games (Bonanno & Kommers, 2005; Bourgonjon et al., 2010; McFarlane, Sparrowhawk, & Heald, 2002) attitudes toward digital games (Bonanno & Kommers, 2008; Bourgonjon et al., 2010), and types of games preferred (Bonanno & Kommers, 2005; Bourgonjon et al., 2010). Based on these differences, some have posited that achievement gains due to digital gaming would be different for boys and girls (Carr & Pelletier, 2008; Dawes & Dumbleton, 2002). However, the results of this study indicated that achievement gains related to treatment (playing *Labyrinth*) were similar for boys and girls.

Furthermore, the pattern of achievement gains seen in this study support the conclusion of Vogel and colleagues' 2006 meta-analysis discussed in Chapter Two. Vogel et al. (2006) reviewed 32 empirical studies and concluded that interactive simulations and games were more effective than traditional classroom instruction on learners' cognitive gains. Similarly, the results of this study indicated that the treatment group who played *Labyrinth* and attended the traditional classrooms achieved higher mathematics scores than the control group who only attended traditional classrooms.

Situational Motivation and *Lure of the Labyrinth*

The Course Interest Survey (CIS) is based on Keller's ARCS model of motivation (1987) and includes four related measures of situational motivation (a) Attention, (b) Relevance, (c) Confidence, and (d) Satisfaction. Model 3 analyzed the effects of treatment on student response on each of these scales while controlling for individual differences by using pretest scores as covariates. Of the four measures of situational motivation used, MANCOVA showed that treatment and control groups only had a significant difference on the Attention scale. The treatment group was the reference group, thus the *b* coefficient of -1.17 ($p < .0005$) showed that students who played *Labyrinth* gained an average of 1.17 points more than those in the control group between pre- and post-test. The significant positive increase on the Attention scale represented a moderate effect size (ES = .42). Keller stated that the attention category refers to those things that gain the learner's attention, build curiosity, and sustain active engagement (Keller, 2008b). Since the CIS scales were designed to measure the fundamental construct of situational motivation a significant change in one scale score can be interpreted as a change in the fundamental construct of situational motivation. Additionally, a shift in situational motivation, particularly the increase in Attention could be responsible for the positive increase in achievement scores. Future research should explore whether situational motivation is a mediating factor for achievement gains.

The increase in attention found in this study was consistent with the findings of Yip and Kwan (2006). Further, these results were consistent with other work that has linked achievement gains through digital game playing with increased motivation (Ju & Wagner, 1997; Kafai, 2001; Kiili, 2005; Lardinois, 1989; Rieber, 1996; Romme, 2003; Rosas et al., 2003). However, Clark proposed that gains seen due to the use of new media were based on a "novelty effect" rather than a true effect of treatment (1983). The nine-week duration of treatment in this study was insufficient exposure to *Labyrinth* to reduce the novelty effect, which could account for the improvement in situational motivation on the Attention scale found in this study. Thus, the findings of this study support the need for longer-term research on the effects of playing digital games to distinguish between treatment and novelty effects.

Model 4B examined whether the effects of playing *Labyrinth* on motivational measures were the same for boys and girls in the treatment and control groups. Statistically the increase in Attention scores did not differ for boys versus girls related to treatment (b = -.51, p = *ns*). This research finding is consistent with the earlier work of Ke (2008) and Papastergiou (2009) who found similar motivation increases for both boys and girls related to digital game use. While not statistically significant, the only group who showed a negative change on the Attention measure from pre- to post-test was the control group girls. Several researchers have noted ongoing concerns about the numbers of girls pursuing advanced learning and careers in mathematics and science (Halpern et al., 2007; NCES, 2004; NSF, 2011). Thus, the significant positive effect on both boys' and girls' Attention related to mathematics from playing *Labyrinth* (b = -1.17, p < .0005) versus a slight, statistically insignificant decline in girls' Attention level in the control group merits further exploration.

Model 4B also demonstrated that the pattern of change for boys and girls in the treatment group did not match the pattern of change for boys and girls in the control group on the Satisfaction measure of motivation (b = 1.15, p < .05). Boys who played *Labyrinth* reported a decrease in their Satisfaction level as compared to boys who did not play the game. Alternately, girls who played *Labyrinth* reported an increase in their Satisfaction level as compared to girls who did not play the game (ES =.09). According to Keller (1987), the Satisfaction category refers to students' positive feelings about their learning experiences, including affirmation that the instructional content was relevant and that they had the ability to learn the material. A possible interpretation of this finding is that boys found the traditional practice classroom sessions (the control group activity while the treatment group engaged in playing *Labyrinth*) a more positive learning experience than their female classmates did. This interpretation is consistent with research findings that as girls move through school they report declining interest in and enjoyment of mathematics (Halpern et al., 2007; NCES, 2004; NSF, 2011). While Robinson and Lubienski (2010) and Stubits et al. (2011) research focused sex differences in mathematics achievement, it may provide some insight as to why girls reported high Satisfaction from playing Labyrinth. These researchers posited that girls experience greater societal pressure to conform to rules, and thus to do math the way the way they were taught. Perhaps playing *Labyrinth* provided an opportunity for girls to creatively experiment with mathematical problem solving, which in turn led to girls developing a more positive perception of their mathematics learning. Thus, the results of this study indicate that further research on sex difference related to mathematics is needed to identify motivational effects in general as well as specifically related to playing serious digital games.

Attitudes Toward Mathematics and *Lure of the Labyrinth*

Model 5 examined changes on students' reporting on five of the subscales of the Fennema-Sherman Mathematics Attitudes Scales (FSMAS): Confidence in Learning Mathematics scale, Attitude toward Success in Mathematics scale, Usefulness of Mathematics scale, Mathematics Anxiety scale, and Effectance Motivation scale. The treatment group showed a significant positive increase on the Usefulness of Mathematics scale ($b = -.77$, $p < .05$, ES = .09) and the Effectance Motivation scale ($b = -1.18$, $p < .0005$, ES = .13). Since all scales of the FSMAS are designed to measure the fundamental construct of attitudes toward mathematics, a change in any scale scores can be interpreted as a change in mathematical attitudes.

Finding that both achievement and attitude toward mathematics were positively affected by treatment (playing Labyrinth) is consistent with research discussed in Chapter Two

(Kadijevich, 2008; Ma & Kishor, 1997; Ma & Xu, 2004; Minato & Kamada, 1996; Singh, Granville, & Dika, 2002; Tocci & Engelhard, 1991). Ma and Kishor's meta-analytic review found an overall positive population correlation of 0.12 between attitudes and achievement in mathematics across studies (1997). Kadijevich's examination of The Trends in International Mathematics and Science Study (TIMMS) 2003 data found significant correlations between attitudes and mathematic achievement scores in 31 out of the 33 participating countries (2008). Kadijevich also separated attitudes into three dimensions, Self-Confidence in Learning Mathematics, Liking Mathematics, and Usefulness of Mathematics. When the data for all 33 participating countries was analyzed, each of the three dimensions outlined by Kadijevich was positively correlated with mathematical achievement. However, in this study only Usefulness and Effectance Motivation scores were significantly different in the treatment group. Though researchers have not yet developed a unified theory to explain how different dimensions of the affective domain relate to mathematical achievement, this study confirms the importance of further research in this area.

The small (ES = .09) but significant ($b = -.77$, $p < .05$) increase on the Usefulness of Mathematics scale for students in the treatment group, indicated a shift in the value participants placed on studying mathematics. Theoretically, this attitude measure should align with the CIS Relevance measure of situational motivation. It is unclear why treatment, i.e. playing *Labyrinth*, resulted in a significant shift on the FSMAS Usefulness scale but not the CIS Relevance scale. However, a possible interpretation is that the FSMAS Usefulness scale is designed to measure an overall perspective of the value of mathematics, whereas the CIS Relevance scale is designed to measure students' perceptions of the connection between the taught information and their lives. The positive effect on the treatment group scores accompanied a drop in the control group's scores on Usefulness. The overall significant finding may be attributable to a negative change in perceptions of the value of studying mathematics from the control group who had more classroom math time. Ultimately, shifts on attitude measure of Usefulness indicate that future research is needed to explore what factors relate to students' perceptions of the value of studying mathematics.

Another measure of students' attitudes toward mathematics, the Effectance Motivation scale, also showed a small but significant increase in the treatment group (b = -1.18, $p < .0005$, ES = .13). White defined effectance motivation as the persistence in learning or exploring a topic or phenomenon for the sole reward of engaging in it (1959). The Effectance Motivation scale was designed to measure effectance in mathematics as a dimension of motivation that ranged from lack of involvement to active enjoyment and seeking challenge in mathematics (Fennema & Sherman, 1976). It suggests that playing a serious digital game, in this case *Labyrinth*, increased students' positive perspective on problem solving in mathematics.

Model 6B examined the joint effect of Sex and Treatment to test whether the pattern of change on attitudes differed for boys and girls in the treatment group versus boys and girls in the control group. The interaction effect was not significant for four of the five measures of attitudes: Confidence in Learning Mathematics, Attitude Toward Success in Mathematics scale, Usefulness of Mathematics scale, and Effectance Motivation scale. Thus, the increases on the Effectance Motivation scale and Usefulness scale were not significantly different for boys than girls.

However, this results revealed a significant treatment and sex interaction effect of a small size (ES = .08) for one of the five attitudes: Anxiety (b = 1.10, $p < .05$). Treatment group girls showed improvement on Anxiety (indicated by increased Anxiety scores), whereas control group girls and treatment group boys showed declines suggesting increased math related anxiety. These results are similar to those found on the CIS situational motivation Satisfaction scale. Girls in the treatment group appear to have found the treatment experience reduced their math anxiety (ES = .08) and increased their sense of having a positive learning environment in their math class (ES = .09). Conversely, girls in the control group showed negative change patterns on both the FSMAS Anxiety and CIS Satisfaction scores. A possible avenue for future research regarding these effects may lie in the "good girl" effect. Forgasz and Leder (2001) proposed that girls are socialized into a "good-girl" role in a classroom environment. Robinson and Lubienski (2010) and Stubits et al. (2011) linked these social messages with the finding that girls are more likely to do math the way they have been told. While these researchers were exploring sex differences in mathematical achievement, it is possible that serious games provide an opportunity for girls to break away from the "good girl" approach and creatively engage with mathematics. This creative play may contribute to both an improved perception of the learning environment as well as a more relaxed sense of self in relation to mathematics.

Limitations

This study examined data collected during a program evaluation that used an experimental design. However, generalization of the results is limited to situations with a similar game and population. Although the sample size was sufficient to support interpretation of the results, the sample was drawn from students at only one international school. Additionally, the cultural diversity of an international school community raises the question about whether cultural background may influence the effects of digital games on students. Replicating the study in multiple international school settings would offer an opportunity to confirm results as well as investigate whether or not differences exist between people's cultural backgrounds and the effects of playing digital games. In addition, the sample group comprised socio-economically privileged students. Future studies are needed to examine whether any connection exists between social or economic status and the effects of serious digital games. Although the experimental design and statistical methods used indicated a causal connection between treatment and positive math outcomes, qualitative inquiry was not used to further validate the results or enrich them.

While the designers of Labyrinth classify it as a serious game, there is insufficient research on the salient factors of different game categories to support generalizing the results of using *Labyrinth* to other serious games. In addition, it is unclear if the effects of treatment were due to playing *Labyrinth*, working on a computer, doing something different from a typical math lesson or some other factor students in the treatment group perceived. Clark suggested that in many cases the learning gains reported in conjunction with the use of new media were due to a "novelty effect" and decreased or disappeared over time (1983). Longer term study of the effects of playing *Labyrinth* would help to determine whether novelty rather than game play was responsible for changes in students' performance, motivation, and attitudes.

The design of the program evaluation did not control for students' awareness that they were participating in a study. Thus, it is possible that the positive results in this study were an experimental artifact due to students' perception of cues or "demand characteristics" which influenced their interpretation of how they were supposed to behave and what outcomes were expected (Orne, 1962). Based on perceived demand characteristics, students in the treatment group may have been inclined to give more positive responses on survey items. In addition, knowledge that they were participating in a program evaluation may have altered students' study patterns resulting in increased achievement scores. While Orne would contend that all studies have demand characteristics, not all demand characteristics interact with the independent variables (1969/2009). Thus, future studies in which groups are matched to elicit similar demand characteristics (e.g. both control and experimental groups are led to believe that they are receiving special treatment) or longer term studies in which the control and experimental groups are geographically isolated and treatment is embedded in classroom routines so is not perceived as a variable of interest are warranted to distinguish between true treatment and Hawthorne effects due to demand characteristics.

Contributions to Research

The learning effectiveness of digital games has been subject to discussion and debate by a number of researchers such as Hays (2005), Mitchell and Savill-Smith (2004), and Girard et al. (2012). One of the criticisms often raised regarding the research base for learning with digital games is the small number of existing empirical studies performed using adequate sample size, a control group for comparison, and equal instructional time between groups (Girard et al., 2012; Harris, 2001; Hays, 2005; Vogel et al., 2006). Thus, the results from this study provide an important contribution to the scholarly discussion of the effectiveness of digital games for learning.

Researchers have extensively examined differences between boys and girls related to both mathematics (Freeman, 2004; Hyde, 2005; Hyde et al. 2008; Robinson & Lubienski, 2010; Stubits et al, 2011) and digital games (Bonanno & Kommers, 2005; Bonanno & Kommers, 2008; Bourgonjon et al, 2010; Cassell & Jenkins, 1998). However, this study's findings related to the different change patterns for boys and girls related to mathematical anxiety and satisfaction based on playing *Labyrinth* versus in-class practice, may provide new insight in these areas.

Increasingly, educational leaders are expected to positively manage the infusion of new technologies into the educational environment, effectively support digitally engaged students (McLeod, 2011; Prensky, 2001), and prepare learners for technologically-rich contexts (Friedman, 2005; McLeod, 2011). For educational leaders one of the main barriers to greater technology integration is the relative lack of research on the efficacy of technology-based instruction (Groff & Mouza, 2008). Thus, the results of this study add to the research base used by educational leaders in making important decisions about access and resource allocation related to technology (Klopfer, et al., 2009).

Contributions to Practice

The analysis of this data set provides further support for integrating digital games into the math classroom. Serious games are a tool of great promise as educators look for methods to improve mathematical achievement, motivation, and attitudes. These methods may include developing new models for achieving instruction differentiation in the classroom setting, seeking tools to eliminate the "seat-time" requirements for students while moving to "performance-based" student accountability, and investigating ways to improve delivery of online lessons. Though women have made tremendous achievements in careers that were once dominated by men, research shows that females are still less likely to pursue math related course work and professions (Halpern et al., 2007; NCES, 2004; NSF, 2011. This study indicated that a serious game such as *Lure of the Labyrinth* may lower girls' anxiety and improve their satisfaction regarding the study of mathematics. If schools incorporate serious games into their math programs, perhaps girls' lower anxiety and improved satisfaction will have long term effects on their academic and career choices, thereby further diminishing the gap between "man's work" and "women's work."

As schools move through the 21st century, educational leaders are increasingly required to effectively oversee the integration of rapidly changing computer technology (ISTE, 2009). "With the world literally at their fingertips, today's students need teachers and administrators to re-envision the role of technology in the classroom" (Blair, 2012, p. 8). However, as school leaders look to break away from obsolete pedagogical practices and effectively integrate technology, the goal remains to implement research-based practices for improved learning (Collins & Halverson, 2009; Halverson, 2005; Halverson, Grigg, Prichett, & Thomas, 2005). In the age of data-driven school reform, the limited and ambiguous research base regarding the effectiveness of digital gaming as an instructional technique presents educational leaders with a barrier against greater adoption (Groff & Mouza, 2008). The results of this study were consistent with other studies indicating digital games show promise for improving student outcomes (Barab, et al., 2007; Vogel et al, 2006), making it incumbent upon educational leaders to consider the role of gaming in education (Collins & Halverson, 2009; Halverson, 2005). Ultimately, "the responsibility for realizing the power of game-based learning environments across schools lies mainly with school leaders" (Halverson, 2005, p. 3).

Conclusion and Future Directions

This study showed that for middle school boys and girls playing *Labyrinth* resulted in a small improvement in mathematics achievement (ES = .13), a moderate increase in situational motivation (ES = .42), and small positive shifts in attitudes (ES = .09, ES = .13). In addition, playing *Labyrinth* lowered girls' reporting of math related anxiety (ES = .08) and increased their self-reported scores on satisfaction with math class (ES = .09). The results of the current study may help teachers, educational leaders, and instructional designers reach firmer conclusions on the effectiveness of digital games.

This study indicated several future avenues of research including exploring whether increased situational motivation and attitudes toward mathematics are correlated or mediating variables for increased achievement. In addition to replicating this study in different settings, it would be valuable to examine the effects of playing *Labyrinth* with different populations. For example, would patterns of change related to digital games be similar for students with special educational needs, students of different cultures, or students of different socio-economic status? Further, the findings of different patterns of change for boys versus girls related to treatment on some measures of motivation and attitudes toward math, indicates that more research is needed to explore sex differences and the effects of playing digital games. In addition, future work should consider assessing multiple forms of treatment to begin to compare the effects of different strategies for improving mathematical learning. Possible areas of interest include comparing the effects on math achievement, motivation and attitudes of some of the following strategies along with in class use of *Labyrinth*: out of class game play, different serious games, drill and practice games, non-digital games, collaboration, and "flipped" instruction. Ultimately, this study's positive findings related to playing *Labyrinth* justify further quantitative and qualitative investigation of the effects of playing serious digital games on students' mathematical achievement, motivation, and attitudes.

In future studies, several methodological issues should be considered. For example, using a digital game, motivation survey, and attitude survey all designed based on the same motivation and learning theories would provide greater understanding of the interaction between game play and changes in the affective domain. In addition, integrating games into normal classroom activities and using an external control group may help reduce the possibility of demand characteristics and a resulting Hawthorne effect. Lastly, longer term or longitudinal exploration of the effects of digital games on students' learning, motivation, and attitudes would provide a more accurate representation of change due to game play and reduce the likelihood of novelty effects.

Lastly, in our era of data driven school reform, many educational leaders recognize a need to focus on research-based improvements to technology integration. As important as empirical evidence may be for educational leaders in justifying decisions about resource allocation, future studies should qualitatively explore any quantitative results to broaden and enrich the understanding of why and how serious games affect students. Through ongoing qualitative and quantitative exploration of the effects of bringing the technology of students' lives into schools, researchers can support the development of educational "leaders who are brave enough to create the new paradigm instead of simply tweaking the status quo and who have the knowledge and ability to create schools that are relevant to the needs of students, families, and society" (McLeod, 2011, p. 4).

References

About NCTM: Mission, vision, and priorities (n.d). Retrieved from http://www.nctm.org/about/content.aspx?id=172

Achieve (2009). *An analysis of the Maryland PreK-8 Voluntary State Curriculum in Mathematics*. Retrieved from http://www.marylandpublicschools.org/NR/rdonlyres/D1CD6035-498E-4E56-B801-1F2A1771FA7D/21093/2009_MD_PreK8_VSC_Math.doc

American Psychological Association (2010). *Publication manual of the American Psychological Association.* Washington, DC: American Psychological Association.

Amory, A., Naicker, K., Vincent, J., & Adams, C. (1999). The use of computer games as an educational tool: identification of appropriate game types and game elements. *British Journal of Educational Technology, 30*, 311-321.

Anderson, C. (2004). An update on the effects of playing violent video games. *Journal of Adolescence, 27*, 113–122.

Anderson, C., & Bushman, B. (2001). Effects of violent video games on aggressive behavior, aggressive cognition, aggressive affect, physiological arousal and prosocial behavior: A meta-analysis. *Psychology and Science, 12*, 353–359.

Anderson, C., & Dill, K. (2000) Video games and aggressive thoughts, feelings and behavior in the laboratory and in life. *Journal of Personality and Social Psychology, 78*, 772–790.

Anderson, J. Q., Boyles, J. L., & Rainie, L. (2012). *The future impact of the Internet on higher education*. Washington, DC: Pew Internet & American Life Project.

Annetta, L. A., Murray, M. R., Laird, S. G., Bohr, S. C., & Park, J. C. (2006). Serious games: Incorporating video games in the classroom. *EDUCAUSE Quarterly, 29*(3), 16-22.

Appleton, J. J., Christenson, S. L., & Furlong, M. J. (2008). Student engagement with school: Critical conceptual and methodological issues of the construct. *Psychology in the Schools, 45*, 369-386.

Australian Council for Educational Research (n.d.a). *The International Schools' Assessment (ISA) construct and content validity.* Retrieved from http://www.acer.edu.au/documents/ISA_ConstructandContentValidity.pdf

Australian Council for Educational Research (n.d.b). *Some technical notes on the International Schools' Assessment.* Retrieved from http://www.acer.edu.au/documents/ISA_ConstructandContentValidity.pdf

Barab, S., Arici, A., & Jackson, C. (2005). Eat your vegetables and do your homework: A design-based investigation of enjoyment and meaning in learning. *Educational Technology*, *16*(1), 15-21.

Barab, S., Dodge, T., Tuzun, H., Job-Sluder, K., Jackson, C., Arici, A., et al. (2007). The Quest Atlantis Project: A socially-responsive play space for learning. In B. E. Shelton & D. Wiley (Eds.), *The Educational Design and Use of Simulation Computer Games* (pp. 159-186). Rotterdam, The Netherlands: Sense Publishers.

Barab, S. A., Gresalfi, M., Dodge, T., & Ingram-Goble, A. (2010). Narratizing disciplines and disciplinizing narratives: Games as 21st century curriculum. *International Journal of Gaming and Computer-Mediated Simulations, 2*(1), 17-30.

Bartholow, B. & Anderson, C. (2002). Effects of violent video games on aggressive behavior: Potential sex differences. *Journal of Experimental Social Psychology, 38*, 283–290.

Bartholow, B., Bushman, B., & Sestir, M. (2006). Chronic violent video game exposure and desensitization to violence: Behavioral and event related brain potential. *Journal of Experimental Social Psychology, 42,* 532–539.

Benbow, C. P., & Stanley, J. C. (1980). Sex differences in mathematical ability: Fact or artifact? *Science, 210,* 1262-1264.

Blair, N. (2012). Technology integration for the new 21st century learner. *Principal*, *91*(3), 8-11.

Block, J. J. (2007). Lessons from Columbine: Virtual and real rage. *American Journal of Forensic Psychiatry, 28*(2), 5-33.

Bonanno, P., & Kommers, P. (2005). Gender differences and styles in the use of digital games. *Educational Psychology, 25*(1), 13–41.

Bonanno, P., & Kommers, P. (2008). Exploring the influence of gender and gaming competence on attitudes towards using instructional games. *British Journal of Educational Technology*, *39*(1), 97–109.

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

Broadbooks, W. J., Elmore, P. B., Pedersen, K., & Bleyer, D. R. (1981). A construct validation study of the Fennema-Sherman Mathematics Attitudes Scales. *Educational and Psychological Measurement, 41,* 551–557.

Campbell, D. T. (1986). Relabeling internal and external validity for applied social scientists. *New Directions for Program Evaluation, 1986*(31), 67-77.

Carr, D., & Pelletier, C. (2008). Games, gender and representation. In E. R. Ferdig (Ed.). *Handbook of Research on Effective Electronic Gaming in Education* (Vol. 2, pp. 911–921). Hershey, PA: Information Science Reference.

Carson, H. C. (2006). The relationship between hypermedia producers' preferred learning styles and the motivational aspects of their productions. *Journal of Education for Library and Information Science, 47*, 106-126.

Cassell, J., & Jenkins, H. (Eds.). (1998). *From Barbie to Mortal Kombat: Gender and computer games*. Cambridge, MA: MIT Press.

Catsambis, S. (2005). The gender gap in mathematics: Merely a step function? In A. M. Gallagher & J. C. Kaufman (Eds.), *Gender differences in mathematics: An integrative psychological approach* (pp. 220-245). Cambridge, UK: Cambridge University Press.

Charles, D., & McAlister, M. (2004). Integrating ideas about invisible playgrounds from play theory into online educational digital games. In M. Rauterberg (Ed.), *Entertainment Computing* (pp. 598-601). Berlin / Heidelberg, Germany: Springer: Lecture Notes in Computer Science

Chatfield, T. (2010) *Fun inc.: Why gaming will dominate the twenty-first century*. New York, NY: Pegasus Books LLC.

Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research*, 53, 445–459.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale. NJ: Erlbaum.

Collins, A. & Halverson, R. (2009*) Rethinking education in the age of technology: The digital revolution and schooling in America*. New York, NY: Teachers College Press.

Csikszentmihalyi, M. (1996). *Creativity: Flow and the Psychology of Discovery and Invention*. New York, NY: Harper Collins.

Darling-Hammond, L. (2009). Teaching and educational transformation. In A. Hargreaves, A.

Lieberman, M. Fullan, & D. Hopkins (Eds.). *Second international handbook of educational change* (pp. 505-520). Netherlands: Springer.

Dawes, L., & Dumbleton, T. (2002). *Computer games in education project*. UK: Becta.

De Jean, J., Upitis, R., Koch, C., & Young, J. (1999). The story of 'Phoenix Quest': How girls respond to a prototype language and mathematics computer game. *Gender and Education, 11*, 207–223.

Denis, G., & Jouvelot, P. (2005). *Motivation-driven educational game design: Applying best practices to music education.* Paper presented at the 2005 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology, Valencia, Spain.

Dewey, J. (1910/1997). *How we think*. Mineola, NY: Dover.

Dickey, M. D. (2005). Engaging by design: How engagement strategies in popular computer and video games can inform instructional design. *Educational Technology Research and Development, 53*(2), 67-83.

Dickey, M. D. (2006). Game design narrative for learning: Appropriating adventure game design narrative devices and techniques for the design of interactive learning environments. *Educational Technology Research and Development, 54*, 245-263.

Dondlinger, M.J. (2007) Educational video game design: A review of the literature. *Journal of Applied Educational Technology, 4*(1), 21-31, 2007.

Eccles, J. S., Wigfield, A., Flanagan, C. A., Miller, C., Reuman, D. A., & Yee, D. (1989). Self concepts, domain values, and self esteem: Relations and changes at early adolescence. *Journal of Personality, 57*, 283-310.

Eccles, J., Wigfield, A., Harold, R. D., & Blumenfeld, P. (1993). Age and gender differences in children's self-and task perceptions during elementary school. *Child Development, 64*, 830-847.

Ellis, H., Heppell, S., Kirriemuir, J., Krotoski, A., & McFarlane, A. (2006). *Unlimited learning: Computer and video games in the learning landscape*. London, UK: Entertainment and Leisure Software Publishers Association.

Ellis, M. (2008). Leaving no child behind yet allowing none too far ahead: Ensuring (in)equity in mathematics education through the science of measurement and instruction. *Teachers College Record*, *110*, 1330–1356.

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin, 136*(1), 101–127.

Facebook (2012). *Key facts.* Retrieved from http://newsroom.fb.com/Key-Facts

Fennema, E., & Sherman, J. A. (1976). Fennema-Sherman mathematics attitudes scales: Instruments designed to measure attitudes toward the learning of mathematics by females and males. *Journal for Research in Mathematics Education, 7*, 324-326.

Ferguson, C. J., Rueda, S. M., Cruz, A. M., Ferguson, D. E., Fritz, S., & Smith, S. M. (2008). Violent video games and aggression. *Criminal Justice and Behavior, 35,* 311-332.

Ferguson, C. J. (2007). The good, the bad and the ugly: A meta-analytic review of positive and negative effects of violent video games. *Psychiatric Quarterly, 78,* 309-316.

Frand, J. L. (2000). The information-age mindset changes in students and implications for higher education. *EDUCAUSE Review, 35*(5), 14-25.

Franklin, S., Peat, M., & Lewis, A. (2003). Non-traditional interventions to stimulate discussion: The use of games and puzzles. *Journal of Biological Education, 37*(2), 79-84.

Freeman, C.E. (2004). Trends in educational equity of girls & women: 2004 (NCES 2005-016). U.S. Department of Education, National Center for Education Statistics. Washington, DC: Government Printing Office.

Friedman T.L. (2005). *The world is flat: A brief history of the twenty-first century*. New York, NY: Farrar, Straus and Giroux.

Furlong, M. J., & Christenson, S. L. (2008). Engaging students at school and with learning: A relevant construct for all students. *Psychology in the Schools, 45*, 365-368.

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

Gee, J. P. (2005). Learning by design: Good video games as learning machines. *E-Learning and Digital Media, 2*(1), 5-16.

Gentile, D. A. (2009). Pathological video game use among youth ages 8 to 18: A national study. *Psychological Science, 20*, 594–602.

Gentile, D. A., Choo, H., Liau, A., Sim, T., Li, D., Fung, D., & Khoo, H. (2011). Pathological video game use among youths: A two-year longitudinal study. *Pediatrics, 127*, 319-329.

Ginn, W. Y. (1995). Jean Piaget: Intellectual development [Electronic Version]. Retrieved from http://www.schoolsalive.com/piaget.pdf.

Ginsburg, A., Cooke, G., Leinwand, S., Noell, J., & Pollock, E. (2005). *International mathematics performance: New findings from the 2003 TIMSS and PISA.* Washington, DC: American Institutes for Research.

Girard, C., Ecalle, J., & Magnan, A. (2012). Serious games as new educational tools: how effective are they? A meta-analysis of recent studies. *Journal of Computer Assisted Learning*. Advance online publication. doi: 10.1111/j.1365-2729.2012.00489.x

Green, H., & Hannon, C. (2007). *Their space education for a digital generation*. Retrieved from http://www.demos.co.uk/publications/theirspace.

Groff, J., & Mouza, C. (2008). A framework for addressing challenges to classroom technology use. *AACE Journal, 16*(1), 21-46.

Grüsser, S. M., Thalemann, R., & Griffiths, M. D. (2007). Excessive computer game playing: Evidence for addiction and aggression? *Cyberpsychology & Behavior, 10,* 290–292.

Hall, S. G. (1912). *Genetic Philosophy of Education.* New York, NY: Sturgis & Walton Company.

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest, 8*(1), 1-51.

Halverson, R. (2005). What can K-12 school leaders learn from video games and gaming*?* *Innovate: Journal of Online Education, 1*(6). Retrieved from http://www.innovateonline.info/index.php?view=article&id=81

Halverson, R., Grigg, J., Prichett, R., & Thomas, C. (2005). *The new instructional leadership: Creating data-driven instructional system in schools* (WCER Working Report No. 2005-9). Madison, WI: Wisconsin Center for Education Research.

Harris, J. (2001). *The effects of computer games on young children – A review of the research* (RDS Occasional Paper No. 72). London, UK: Research, Development and Statistics Directorate, Communications Development Unit, Home Office.

Hays, R. T. (2005). *The effectiveness of instructional games: A literature review and discussion*. Retrieved from Naval Air Warfare Center Training System Division website: http://stinet.dtic.mil/oai/oai?&verb=getRecord&metadataPrefix=html&ide

Holden, C. (2010). Behavioral addictions debut in proposed DSM-V. *Science, 327*, 935.

Holland, W., Jenkins, H., & Squire, K. (2002). Game theory. In B. Perron & M. Wolf (Eds.), *Video game theory*. Retrieved from http://www.educationarcade.org/gtt/.

Howe, N., & Strauss, W. (2000). *Millennials rising: The next great generation*. New York, NY: Vintage Books.

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist, 60,* 581-592.

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performance. *Science, 321,* 494-495.

The International Society for Technology in Education [ISTE] (2009). *The ISTE NETS and performance indicators for administrators (NETS•A).* Retrieved from http://www.iste.org/standards/nets-for-administrators.aspx

Internet World Stats (2012, June 30). Retrieved from http://www.internetworldstats.com/stats.htm

Jackson, C. D. & Leffingwell, R. J. (1999). The role of instructions in creating math anxiety in students from kindergarten through college. *The Mathematics Teacher, 92,* 583-586.

James, W. (1890/2007). *The principles of psychology* (Vol. 2). New York, NY: Cosimo Inc.

Jones, G., & Kalinowski, K. (2007). Touring Mars online, real-time, in 3D for mathematics and science educators and students. *Journal of Computers in Mathematics and Science Teaching, 26*(2), 123-136.

Ju, E., & Wagner, C. (1997). Personal computer adventure games: their structure, principles, and applicability for training. *ACM SIGMIS Database, 28*(2), 78-92.

Kadijevich, D. (2008). TIMMS 2003: Relating dimensions of mathematics attitude to mathematics achievement. *Institute for Educational Research, 40,* 257-273.

Kafai, Y. B. (2001). *The educational potential of electronic games: From games--to--teach to games--to—learn.* Chicago, IL: Cultural Policy Center, University of Chicago.

Ke, F. (2008). A case study of computer gaming for math: Engaged learning from gameplay? *Computers & Education*, *51*, 1609-1620.

Ke, F., & Grabowski, B. (2007). Gameplaying for maths learning: Cooperative or not? *British Journal of Educational Technology, 38*(2), 249-259.

Kebritchi, M., Hirumi, A., & Bai, H. (2010). The effects of modern mathematics computer games on mathematics achievement and class motivation. *Computers & Education, 55*(2), 427-443.

Keller, J. M. (1987). Development and use of the ARCS model of instructional design. *Journal of Instructional Development*, *10*(3), 2-10.

Keller, J. M. (2008). First principles of motivation to learn and e^3-learning. *Distance Education, 29*(2), 175-186.

Keller, J. M. (2008). An integrative theory of motivation, volition, and performance. *Technology, Instruction, Cognition, and Learning, 6*(2), 79-104.

Keller, J. M. (2010). *Motivational design for learning and performance: The ARCS model approach*. New York, NY: Springer Science and Business Media.

Keller, J. M., & Suzuki, K. (2004). Learner motivation and E-learning design: A multinationally validated process. *Journal of Educational Media, 29*, 229-239.

Kiili, K. (2005). Content creation challenges and flow experiences in educational games: The IT-Emperor case. *The Internet and Higher Education, 8*(3), 183-198.

Kilpatrick, J. (2003). Introduction. In J. Kilpatrick, W.G. Martin & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 1-4). Reston, VA: The National Council of Teachers of Mathematics.

Kim, C., & Keller, J. M. (2010). Motivation, volition and belief change strategies to improve mathematics learning. *Journal of Computer Assisted Learning, 26,* 407-420.

Kirriemuir, J., & McFarlane, A. (2004). *Literature review in games and learning* (Report No. 8). England, UK: Futurelab.

Klopfer, E., Osterweil, S., Groff, J., Haas, J., (2009). Using the technology of today, in the classroom today: The instructional power of digital games, social networking, and simulations, and how teachers can leverage them. Retrieved from The Education Arcade website:
http://education.mit.edu/papers/GamesSimsSocNets_EdArcade.pdf

Ko, C. H., Yen, J. Y., Yen, C. F., Lin, H. C., & Yang, M. J. (2007). Factors predictive for incidence and remission of Internet addiction in young adolescents: A prospective study. *Cyberpsychology & Behavior, 10*, 545–555.

Koether, M. C. (2003). The name game: Learning the connectivity between the concepts. *Journal of Chemical Education, 80,* 421-423.

Koster, R. (2004). *A theory of fun for game design.* Scottsdale, AZ: Paraglyph Press Inc.

LaFraniere, S. (2005, August 25). Cellphones catapult rural Africa to 21st century. *The New York Times.* Retrieved from http://www.nytimes.com/2005/08/25/international/africa/25africa.html?pagewanted=1

Lardinois, C. (1989). Simulation, gaming and training in a competitive, multimodal, multicompany, intercity passenger-transportation environment. *The Journal of the Operational Research Society, 40*, 849-861.

Lauer, T. E. (2003). Conceptualizing ecology: a learning cycle approach. *The American Biology Teacher, 65*, 518-522.

Lee, J., Luchini, K., Michael, B., Norris, C., & Soloway, E. (2004). More than just fun and games: Assessing the value of educational video games in the classroom. *CHI 04 Extended Abstracts on Human Factors in Computing Systems*, 1375-1378.

Leinhardt, G., Zaslavsky, O., & Stein, M. K. (1990). Functions, graphs, and graphing: Tasks, learning, and teaching. *Review of Educational Research, 60*(1), 1-64.

Lenhart, A., Kahne, J., Middaugh, E., Macgill, A., Evans, C., & Vitak, J. (2008). *Teens, video games, and civics.* Washington, DC: Pew Internet & American Life Project.

Levin, D., & Arafeh, S. (2003). The Digital Disconnect: The widening gap between Internet-savvy students and their schools. *Technology and Teacher Education Annual, 2,* 1002-1007.

Lure of the Labyrinth: About Us. (n.d). Retrieved from http://labyrinth.thinkport.org/www/about.php

Ma, X., & Kishor, N. (1997). Assessing the relationship between attitude toward mathematics and achievement in mathematics: A meta-analysis. *Journal for Research in Mathematics Education*, *28*(1), 26-47.

Ma, X., & Xu, J. (2004). Determining the causal ordering between attitude toward mathematics and achievement in mathematics. *American Journal of Education*, *110*, 256-280.

Maccoby, E. E., & Jacklin, C. N. (1974). Myth, reality and shades of gray: What we know and don't know about sex differences. *Psychology Today, 8*(7), 109-112.

Marsh T. (2011). Serious games continuum: Between games for purpose and experiential environments for purpose. *Entertainment Computing, 2,* 61–68.

McFarlane, A., Sparrowhawk, A., & Heald, Y. (2002). *Report on the educational use of games*. Retrieved from TEEM: Teachers Evaluating Educational Multimedia website: http://www.teem.org.uk/publications/teem_gamesined_full.pdf

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

McLeod, S. (2011, Summer). Are we irrelevant to the digital, global world in which we now live? *University Council for Educational Administration Review, 52*(2), 1–5.

Michael, D. R., & Chen, S. L. (2006). *Serious games: Games that educate, train, and inform*. Boston, MA: Thomson Course Technology PTR.

Middleton, J. A., & Spanias, P. A. (1999). Motivation for achievement in mathematics: Findings, generalizations, and criticisms of the research. *Journal for Research in Mathematics Education, 30*(1), 65-88.

Minato, S., & Kamada, T. (1996). Results of research studies on causal predominance between achievement and attitude in junior high school mathematics of Japan. *Journal for Research in Mathematics Education, 27*(1), 96-99.

Mitchell, A., & Savill-Smith, C. (2004). *The use of computer and video games for learning: A review of the literature*. London, UK: Learning and Skills Development Agency.

Mulhern, F., & Rae, G. (1998). Development of a shortened form of the Fennema-Sherman Mathematics Attitudes Scales. *Educational and Psychological Measurement. 58,* 295-306.

Nathan, M. J., & Koellner, K. (2007). A framework for understanding and cultivating the transition from arithmetic to algebraic reasoning. *Mathematical Thinking and Learning: An International Journal, 9*(3), 179-192.

National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics*. Reston, VA: NCTM.

National Council of Teachers of Mathematics. (2006). *Curriculum focal points for prekindergarten through grade 8 mathematics: A quest for coherence.* Reston, VA: NCTM.

National Science Foundation, Division of Science Resources Statistics (2011). *Women, Minorities, and Persons with Disabilities in Science and Engineering: 2011* (Special Report NSF 11-309). Retrieved from http://www.nsf.gov/statistics/wmpd/.

Norusis, M. J. (1999). *SPSS 9.0 Guide to data analysis*. Upper Saddle River, NJ: Prentice Hall, Inc.

Nunnaly, J. (1978). *Psychometric theory*. New York: McGraw-Hill.

Oblinger, D. G. (2004). The next generation of educational engagement. *Journal of Interactive Media in Education, 8*(8), 2-19.

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, 776-783.

Orne, M. T. (1969/2009). Demand characteristics and the concept of quasi-controls. In R. Rosenthal and R. Rosnow (Eds.), *Artifacts in behavioral research: Robert Rosenthal and Ralph L. Rosnow's Classic Books* (pp. 110-137). New York: Oxford University Press, Inc.

Ogunlesi, T. & Busari, S. (2012, September 14). Seven ways mobile phones have changed lives in Africa. *CNN*. Retrieved from http://www.cnn.com

Pajares, F. (2005). Gender differences in mathematics self-efficacy beliefs. In A. M. Gallagher & J. C. Kaufman (Eds.), *Gender differences in mathematics: An integrative psychological approach* (pp. 294-312). Cambridge, UK: Cambridge University Press.

Pange, J. (2003). Teaching probabilities and statistics to preschool children. *Information Technology in Childhood Education Annual*, 163-172.

Papastergiou, M. (2009). Digital game-based learning in high school computer science education: Impact on educational effectiveness and student motivation. *Computers &Education*, *52*(1), 1–12.

Pareto, L., Arvemo, T., Dahl, Y., Haake, M. and Gulz, A. (2011). A teachable-agent arithmetic game's effects on mathematics understanding, attitude and self-efficacy. In G. Biswas, S. Bull, J. Kay, & A. Mitrovic (Eds.), *Lecture Notes in Computer Science: Vol. 6738. Artificial Intelligence in Education* (pp. 247-255). doi: 10.1007/978-3-642-21869-9_33

Peng, L.H., & Li, X. (2009). A survey of Chinese college students addicted to video games. *China Education Innovation Herald*, *28*, 111–112.

Perry, A. (2011, June 30). Silicon savanna: Mobile phones transform Africa. *Time.* Retrieved from http://www.time.com/time/magazine/article/0,9171,2080702-2,00.html

Perry, E. L., & Ballou, D. J. (1997). The role of work, play, and fun in microcomputer software training. *ACM SIGMIS Database, 28*(2), 93-112.

Phillips, D. C. (1995). The good, the bad, and the ugly: The many faces of Constructivism. *Educational Research, 24*(7), 5-12.

Pilli, O. & Aksu, M. (2012) The effects of computer-assisted instruction on the achievement, attitudes and retention of fourth grade mathematics students in North Cyprus. *Computers & Education*. Advance online publication. doi: 10.1016/j.compedu.2012.10.010.

Pink, D. (2009). *Drive: The surprising truth about what motivates us.* New York, NY: Riverhead.

Pink, D. (2009, October 18). Games, not grades! [Web log post]. Retrieved from http://www.danpink.com/archives/2009/10/games-not-grades

Porter, G., Starcevic, V., Berle, D., & Fenech, P. (2010). Recognizing problem video game use. *Australian and New Zealand Journal of Psychiatry, 44*, 120–128.

Prensky, M. (2001). Digital natives, digital immigrants Part 1. *On the Horizon, 9*(5), 1-6.

Prensky, M. (2006). *Don't bother me mom—I'm learning!* St. Paul, MN: Paragon House.

Purcell, K., Rainie, L., Rosenstiel, T., & Mitchell, A. (2011) How mobile devices are changing community information environments. Retrieved from http://stateofthemedia.org/2011/Mobile-survey/

Reys, B.J. (2008). *Mathematics curriculum: A vehicle for school improvement*. Retrieved from http://tsg.icme11.org/document/get/350

Rieber, L. P. (1996). Seriously considering play: Designing interactive learning environments based on the blending of microworlds, simulations, and games. *Educational Technology Research and Development, 44*(2), 43-58.

Robinson, J. P., & Lubienski, S. T. (2011). The development of gender achievement gaps in mathematics and reading during elementary and middle school. *American Educational Research Journal, 48*, 268-302.

Romme, A. G. L. (2003). Learning outcomes of microworlds for management education. *Management Learning, 34*(1), 51-61.

Rosas, R., Nussbaum, M., Cumsille, P., Marianov, V., Correa, M., Flores, P., et al. (2003). Beyond Nintendo: Design and assessment of educational video games for first and second grade students. *Computers & Education, 40*(1), 71-24.

Salen, K., & Zimmerman, E. (2004). *Rules of play: Game design fundamentals*. Cambridge, MA: The MIT Press.

Shaffer, D. W. (1997). Learning mathematics through design: The anatomy of Escher's World. *The Journal of Mathematical Behavior, 16*(2), 95-112.

Shaffer, D. W., Squire, K. R., Halverson, R., & Gee, J. P. (2005). Video games and the future of learning. *Phi Delta Kappan, 87*(2), 104-111.

Shernoff, D. J., Csikszentmihalyi, M., Schneider, B., & Shernoff, E. S. (2003). Student engagement in high school classrooms from the perspective of flow theory. *School Psychology Quarterly, 18*(2), 158-176.

Sherry, J. (2001). The effects of violent video games on aggression: A meta-analysis. *Human Communication Research, 27*, 409–431.

Sherry, J. (2008) Violent video games and aggression: Why can't we find links? In R. Preiss, B. Gayle, N. Burrell, M. Allen M, & J. Bryant (Eds.), *Mass media effects research: Advances through meta-analysis* (pp. 231-248). Mahwah, NJ: L. Erlbaum.

Schrenko, L. (1994). *Structuring a learner-centered school.* Arlington Heights, IL: IRI Skylight.

Simpson, E. S. (2005). What teachers need to know about the videogame generation. *TechTrends, 49*(5), 17-22.

Singh, K., Granville, M., & Dika, S. (2002). Mathematics and science achievement: Effects of motivation, interest, and academic engagement. *The Journal of Educational Research, 95*, 323-332.

Sitzmann T. (2011). A meta-analytic examination of the instructional effectiveness of computer-based simulation games. *Personnel Psychology, 64*, 489–528.

Small, R. V., Zakaria, N., & El-Figuigui, H. (2004). Motivational aspects of information literacy skills instruction in community college libraries. *College & Research Libraries, 65*(2), 96-121.

Squire, K. D. (2003). Video games in education. *International Journal of Intelligent Games & Simulation, 2*(1), 49-62.

Squire, K., Giovanetto, L., Devane, B., & Durga, S. (2005). From users to designers: Building a self-organizing game-based learning environment. *TechTrends, 49*(5), 34-42.

Squire, K., & Jenkins, H. (2003). Harnessing the power of games in education. *Insight, 3*(1), 5-33.

Stubits, L., Zackon, J., Roberts, L., Siegal, B., & Flanagan, N. (2011). *Gender differences in math achievement in the early elementary years: Are girls too good?* Paper presented at the meeting of American Educational Research Association, New Orleans, LA.

Sweetser, P., & Wyeth, P. (2005). GameFlow: a model for evaluating player enjoyment in games. *Computers in Entertainment (CIE), 3*(3), 1-24.

Tapscott, D. (1998). *Growing up digital: the rise of the net generation*. New York, NY: McGraw-Hill Companies.

Tapscott, D. (1999). Educating the Net Generation. *Educational Leadership, 56*(5), 6-11.

The Center for the Digital Future (2012). *Special report: America at the digital turning point.* USC Annenberg School. Retrieved from http://www.digitalcenter.org/

Tocci, C. M., & Engelhard, G. (1991). Achievement, parental support, and gender differences in attitudes toward mathematics. *The Journal of Educational Research, 84*, 280-286.

Utsumi, M. C., & Mendes, C. R. (2000). Researching the attitudes towards mathematics in basic education. *Educational Psychology, 20*, 237-243.

Vallerand, R. J., & Ratelle, C. F. (2002). Intrinsic and extrinsic motivation: A hierarchical model. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research.* Rochester, NY: University of Rochester Press.

Van de Walle, J. A., Karp, K. S., & Bay-Williams, J. M. (2010). *Elementary and middle school mathematics: Teaching developmentally* (7th ed.). Boston, MA: Pearson.

Van Eck, R. (2006) Digital game-based learning: It's not just the digital natives who are restless. *EDUCAUSE Review*, *41*(2), 16-30.

Vogel, J. J., Vogel, D. S., Cannon-Bowers, J., Bowers, C. A., Muse, K., & Wright, M. (2006). Computer gaming and interactive simulations for learning: A meta-analysis. *Journal of Educational Computing Research, 34*(3), 229-243.

Volet, S. E. (1997). Cognitive and affective variables in academic learning: the significance of direction and effort in students' goals. *Learning and Instruction, 7,* 235-254.

Vroom, V. H. (1964). *Work and motivation.* New York, NY: Wiley.

Vygotsky, L. (1967). Play and its role in the mental development of the child. *Soviet Psychology*, 5(3), 6-18.

Vygotsky, L. (1978). *Minds in society: the development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Wallis, C., Steptoe, S., & Miranda, C. (2006). How to bring our schools out of the 20th century. *Time*, *168*, 50-56. Retrieved from http://www.time.com/time/magazine/article/0,9171,1568480,00.html.

Waraich, A., Sharman, A., & Mitchell, W. L. (2004). Narrative and meaning in the design of interactive learning environments. *International Journal of Continuing Engineering Education and Life Long Learning, 14*, 472-483.

Weigman, O., & Van Schie, E. (1998) Video game playing and its relations with aggressive and prosocial behavior. *British Journal of Social Psychology, 37*, 367-378.

White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review, 66*, 297-333.

White-Clark, R., DiCarlo, M., & Gilchriest, N. (2008). "Guide on the side": An instructional approach to meet mathematics standards. *The High School Journal, 91*(4), 40-44.

Wigfield, A. (1994). Expectancy-value theory of achievement motivation: A developmental perspective. *Educational Psychology Review, 6*(1), 49-78.

Wigfield, A., Eccles, J. S., Harold, R. D., Arbreton, A. J. A., Freedman-Doan, C., & Blumenfeld, P. C. (1997). Change in children's competence beliefs and subjective task values across the elementary school years: A 3-year study. *Journal of Educational Psychology, 89*, 451-469.

Wigfield, A., Eccles, J. S., Mac Iver, D., Reuman, D. A., & Midgley, C. (1991). Transitions during early adolescence: Changes in children's domain-specific self-perceptions and general self-esteem across the transition to junior high school. *Developmental Psychology, 27*, 552-565.

Wilhelm, T., Carmen, D., & Reynolds, M. (2002). Connecting kids to technology: Challenges and opportunities. Retrieved from Annie E. Casey Foundation website http://www.aecf.org/KnowledgeCenter/Publications.aspx?pubguid={E10D9A20-D2E3-4238-9E29-624C0140DC6C}

Williams, D. (2005). Bridging the methodological divide in game research. *Simulation & Gaming*, *36*, 447-462.

Williams, D., & Skoric, M. (2005). Internet fantasy violence: A test of aggression in an online game. *Communication Monographs*, *72*, 217–233.

Wolf, F. M., & Blixt, S. L. (1981). A cross-sectional cross-lagged panel analysis of mathematics achievement and attitudes: Implications for the interpretation of the direction of predictive validity. *Educational and Psychological measurement*, *41*, 829-834.

Yip, F. W. M., & Kwan, A. C. M. (2006). Online vocabulary games as a tool for teaching and learning English vocabulary. *Educational Media International*, *43*, 233-249.

Zyda, M. (2005) From visual simulation to virtual reality to games. *Computer*, *38*(9), 25-32.

Appendix A

Lure of the Labyrinth Wing One Mathematical Standards

Lure of the Labyrinth (*Labyrinth*) is correlated with the Maryland voluntary state curriculum (VSC) standards and the National Council of Teachers of Mathematics (NCTM) standards. Table A1 shows the Maryland VSC, NCTM, and International School of Kenya (ISK) standards aligned with the topical math content covered in Wing One of *Labyrinth*.

Table A1

Math Topic Covered in Labyrinth and Maryland VSC, NCTM, and ISK related Standards (Strands) and Objectives (Expectations, Benchmarks) for Grade 6-8 Mathematics

Math Topic: *Labyrinth* Wing One	Maryland VSC **Standard** • Grades 6-8 Objectives	NCTM **Standard** • Grades 6-8 Expectations	ISK **Strand** • Grades 6-8 Benchmarks
Equivalent ratios and proportional reasoning	**Knowledge of Number Relationships and Computation/ Arithmetic** • Determine Equivalent Ratios • Solve problems using proportional reasoning • Represent ratios in a variety of forms • Determine and use rates, unit rates, and percents as ratios in the context of a problem • Determine equivalent forms of rational numbers expressed as fractions, decimals, percents, and	**Number and Operations** • Understand and use ratios and proportions to represent quantitative relationships; • Develop meaning for integers and represent and compare quantities with them.	**Number and Operations** • Use unit rate, percents, and proportion to solve problems • Solve and verify practical problems using rational numbers, percents, integers, fractions, proportions and decimals

Math Topic: *Labyrinth* Wing One	Maryland VSC **Standard** • Grades 6-8 Objectives	NCTM **Standard** • Grades 6-8 Expectations	ISK **Strand** • Grades 6-8 Benchmarks
	ratios		
Fractions: adding, subtracting, multiplying, dividing, and equivalence	• Add, subtract and multiply positive fractions and mixed numbers • Determine equivalent forms of rational numbers expressed as fractions, decimals, percents, and ratios	• Work flexibly with fractions, decimals, and percents to solve problems	• Identify equivalence relationships among fractions, decimals, and percents • Solve problems in all four operations that involve integers, factions and decimals • Demonstrate mastery of equivalent fractions and proportions
Algebraic Expressions	**Knowledge of Algebra, Patterns, and Functions** • Write an algebraic expression to represent unknown quantities • Evaluate algebraic expressions	**Algebra** • Represent, analyze, and generalize a variety of patterns with tables, graphs, words, and, when possible, symbolic rules • Relate and compare different forms of representation for a relationship;	**Algebra** • Identify patterns concretely and pictorially • Translate phrases describing simple mathematical relationships into algebraic expressions • Use equations to describe multi-step problems • Apply algebraic concepts to solve and verify practical problems

Appendix B

Modified Course Interest Survey (Keller, 2010)

Student Instructions: There are 34 statements in this section. Please think about each statement in relation to the math class you are taking and indicate how true it is. Give the answer that truly applies to you, and not what you would like to be true, or what you think others want to hear.

Think about each statement by itself and indicate how true it is. Do not be influenced by your answers to other statements.

Click on the circle next to the response that best fits your experience so far. Use the following values to indicate your response to each item.

1 = Not true   2 = Slightly true   3 = Moderately true 4 = Mostly true   5 = Very true

1. My math teacher knows how to make us feel enthusiastic about math.
2. The things I am learning in math class will be useful to me.
3. I feel confident that I will do well in math class.
4. Math class has very little in it that captures my attention.
5. My teacher makes math seem important.
6. You have to be lucky to get good grades in my math class.
7. I have to work too hard to succeed in math class.
8. I do NOT see how the content of my math class relates to anything I already know.
9. Whether or not I succeed in math class is up to me.
10. My math teacher creates suspense when building up to a point.
11. The subject matter of my math class is just too difficult for me.
12. I feel that my math class gives me a lot of satisfaction.
13. In my math class, I try to set and achieve high standards of excellence.
14. I feel the grades or other recognition I receive in math are fair compared to other students.
15. The students in my math class seem curious about the subject matter.
16. I enjoy working for my math class.
17. It is difficult to predict what grade my teacher will give my math assignments
18. I am pleased with my math teacher's evaluations of my work compared to how well I think I have done.

19. I feel satisfied with what I am getting from my math class.
20. The content of math class relates to my expectations and goals.
21. My math teacher does unusual or surprising things that are interesting.
22. The students actively participate in my math class.
23. To accomplish my goals, it is important that I do well in math class.
24. My math teacher uses a variety of teaching techniques.
25. I do NOT think I will benefit much from my math class.
26. I often daydream while in math class.
27. As I am taking this math class, I believe that I can succeed if I try hard enough.
28. The personal benefits of math class are clear to me.
29. My curiosity is often stimulated by the questions asked or the problems given in math class.
30. I find the challenge level in math class to be about right: neither too easy nor too hard.
31. I feel disappointed with math class.
32. I feel that I get enough recognition of my work in math class by means of grades, comments, or other feedback.
33. The amount of work I have to do is appropriate for this type of math class.
34. I get enough feedback to know how well I am doing in math class.

The scoring system for the CIS is shown in Table B1. Table B2 indicates the items associated with each of the four CIS scales. Reverse loaded items were reverse scored. Therefore, a higher score indicated a more positive level of situational motivation on the CIS.

Table B1

Scoring of Course Interest Survey

Score	Indicator Statement	
	Regular Loaded	Reverse Loaded
1	Not true	Very true
2	Slightly true	Mostly true
3	Moderately true	Moderately true
4	Mostly true	Slightly true
5	Very true	Not true

Table B2

Items in CIS Scales

Attention	Relevance	Confidence	Satisfaction
1	2	3	7 (reverse)
4 (reverse)	5	6 (reverse)	12
10	8 (reverse)	9	14
15	13	11 (reverse)	16
21	20	17 (reverse)	18
24	22	27	19
26 (reverse)	23	30	31 (reverse)
29	25 (reverse)	34	32
	28		33

Appendix C

The Fennema-Sherman Mathematics Attitude Scale

Student Instructions: There are 60 statements in this section. Please think about each statement and decide how strongly you agree or disagree with it. Give the answer that truly applies to you, and not what you would like to be true, or what you think others want to hear.

Click on the circle next to the response that best fits your experience so far. Use the following values to indicate your response to each item.

1 = I strongly disagree
2 = I disagree
3 = I'm not sure
4 = I agree
5 = I strongly agree

1. It would make me happy to be recognized as an excellent student in mathematics.
2. I'd be proud to be the outstanding student in mathematics.
3. I am happy to get top grades in mathematics.
4. It would be really great to win a prize in mathematics.
5. Being first in a mathematics competition would make me pleased.
6. Being regarded as smart in mathematics would be a great thing.
7. Winning a prize in mathematics would make me feel unpleasantly conspicuous.
8. People would think I was strange if I got high grades in mathematics.
9. If I got good grades in mathematics, I would try to hide it.
10. If I got the highest grade in mathematics, I'd prefer no one knew.
11. It would make people like me less if I were a really good mathematics student.
12. I don't like people to think I'm smart in mathematics.
13. Generally I have felt secure about attempting mathematics.
14. I am sure I could do advanced work in mathematics.
15. I am sure that I can learn mathematics.

16. I think I could handle more difficult mathematics.

17. I can get good grades in mathematics.

18. I have a lot of self-confidence when it comes to mathematics.

19. I’m no good at mathematics.

20. I don’t think I could do advanced mathematics.

21. I’m not the type to do well in mathematics.

22. For some reason even though I study, mathematics seems unusually hard for me.

23. Most subjects I can handle OK, but I have a knack of messing up in mathematics.

24. Mathematics has been my worst subject.

25. I’ll need mathematics for my future work.

26. I study mathematics because I know how useful it is.

27. Knowing mathematics will help me earn a living.

28. Mathematics is a worthwhile and necessary subject.

29. I’ll need a firm mastery of mathematics for my future work.

30. I will use mathematics in many ways as an adult.

31. Mathematics is of no relevance to my life.

32. Mathematics will not be important to me in my life’s work.

33. I see mathematics as a subject I will rarely use in daily life as an adult.

34. Taking mathematics is a waste of time.

35. In terms of my adult life, it is not important for me to do well in mathematics.

36. I expect to have little use for mathematics when I get out of school.

37. I like mathematics puzzles.

38. Mathematics is enjoyable and stimulating to me.

39. When a mathematics problem arises that I can't immediately solve, I stick with it until I have the solution.
40. Once I start trying to work on a mathematics puzzle I find it hard to stop.
41. When a question is left unanswered in mathematics class, I continue to think about it afterwards.
42. I am challenged by mathematics problems I can't understand immediately.
43. Figuring out mathematics problems does not appeal to me.
44. The challenge of mathematics problems does not appeal to me.
45. Mathematics puzzles are boring.
46. I don't understand how some people can spend so much time on mathematics and seem to enjoy it.
47. I would rather have someone give me a solution to a difficult mathematics problem than to have it work it out for myself.
48. I do as little work in mathematics as possible.
49. Math does not scare me at all.
50. It wouldn't bother me at all to take more math courses.
51. I don't usually worry about being able to solve math problems.
52. I almost never get nervous during a math test.
53. I am usually calm during math tests.
54. I am usually calm in math class.
55. Math usually makes me feel uncomfortable and nervous.
56. Math makes me feel uncomfortable, restless, irritable, and impatient.
57. I get a sick feeling when I think of trying to do math problems.

58. My mind goes blank and I am unable to think clearly when working math problems.

59. A math test would scare me.

60. Math makes me feel uneasy, confused, and nervous.

The scoring for regularly and reverse loaded items is show in Table C1. Table C2 indicates the items associated with each of the five FSMAS scales: Confidence in Learning Mathematics; Attitude toward Success in Mathematics scale; Usefulness of Mathematics scale; Mathematics Anxiety scale; and Effectance Motivation scale. Reverse loaded items were reverse scored. All Math Anxiety Scale items were reverse scored. Therefore, a higher score indicated a more positive attitude toward mathematics on the FSMAS including a lower level of reported math anxiety.

Table C1

Scoring of the Fennema-Sherman Mathematics Attitude Scale

Score	Indicator Statement	
	Regular Loaded	Reverse Loaded
1	I strongly disagree	I strongly agree
2	I disagree	I agree
3	I’m not sure	I’m not sure
4	I agree	I disagree
5	I strongly agree	I strongly disagree

Table C2

Items in the FSMAS Scales

Success	Confidence	Usefulness	Effectance	Anxiety
1-6	13-18	25-30	37-42 (reverse)	49-54
7-12 (reverse)	19-24 (reverse)	31-36 (reverse)	43-48	55-60 (reverse)

Appendix D

Statistical Analysis of Assumptions for MANCOVA

In order for MANCOVA to be an appropriate method of statistical analysis, the data must conform to certain expectations including: normal distribution, homogeneity of variances, homogeneity of covariance, and correlation of dependent variables among themselves. Table D1 shows the results for the tests of each of underlying assumption for MANCOVA. Levene's Test addressed homogeneity of variance, Box's M addressed the homogeneity of covariance for multivariate models (Models 3-6B), Bartlett's Test of Sphericity indicated correlation of dependent variables among themselves for multivariate models (Models 3-6B), and Q-Q Plots of residuals indicated the shape of the distribution of scores.

Table D1

Results of Tests of Assumptions for Each Model

Model	Dependent Variable	*F* for Levene's Test	Box's M	Bartlett's Test of Sphericity	Pretest Equal	Q-Q Plot	Conclusion Regarding Assumptions
1	Achievement	.57	NA	NA	Yes	Normal	Met
2A	Achievement	1.02	NA	NA	Yes	Normal	Met
2B	Achievement	1.20	NA	NA	Yes	Normal	Met
2C	Achievement	.54	NA	NA	Yes	Normal	Met
2D	Achievement	.58	NA	NA	Yes	Normal	Met

3	Attention	2.27	6.46	54.40***	Yes	Neg. skew	Not met[a]
	Relevance	1.21	6.46	54.40***	Yes	Normal	Met
	Confidence	3.23+	6.46	54.40***	Yes	Normal	Met
	Satisfaction	.02	6.46	54.40***	Yes	Pos. skew	Not met[a]
4A	Attention	1.26	19.00	57.09***	Yes	Normal	Met
	Relevance	1.57	19.00	57.09***	Yes	Neg. skew	Not met[a]
	Confidence	2.03	19.00	57.09***	Yes	Normal	Met
	Satisfaction	1.58	19.00	57.09***	Yes	Pos. skew	Not met[a]
4B	Attention	1.32	19.00	57.09***	Yes	Neg. skew	Not met[a]
	Relevance	1.58	19.00	57.09***	Yes	Neg. skew	Not met[a]
	Confidence	1.93	19.00	57.09***	Yes	Normal	Met
	Satisfaction	1.43	19.00	57.09***	Yes	Pos. skew	Not met[a]

5	Confidence	8.57**	16.04	45.22***	Yes	Pos. skew	Not met[a]
	Success	.06	16.04	45.22***	Yes	Pos. skew	Not met[a]
	Usefulness	1.71	16.04	45.22***	Yes	Normal	Met
	Anxiety	.01	16.04	45.22***	Yes	Pos. skew	Not met[a]
	Effectance	5.35	16.04	45.22***	Yes	Normal	Met
6A	Confidence	3.06*	45.73	44.96***	Yes	Normal	Met
	Success	2.03	45.73	44.96***	Yes	Pos. skew	Not met[a]
	Usefulness	1.12	45.73	44.96***	Yes	Normal	Met
	Anxiety	.34	45.73	44.96***	Yes	Pos. skew	Not met[a]
	Effectance	1.98	45.73	44.96***	Yes	Normal	Met
6B	Confidence	3.03*	45.73	47.50***	Yes	Normal	Met
	Success	2.05	45.73	47.50***	Yes	Pos skew	Not met[a]
	Usefulness	1.17	45.73	47.50***	Yes	Normal	Met
	Anxiety	.71	45.73	47.50***	Yes	Pos skew	Not met[a]
	Effectance	1.88	45.73	47.50***	Yes	Normal	Met

[a]Robust to violation

+*p* < .10. **p* < .05. ***p* < .005. ****p* < .0005

Appendix E

Means and Standard Deviations for Relevance, Confidence, and Satisfaction Scores on the CIS

Table E1

Means and Standard Deviations for CIS Relevance Scores

Sex	Group		CIS Relevance pre	CIS Relevance post
Male	Control	Mean	24.95	25.12
		N	42	42
		Std. Deviation	4.345	4.295
	Treatment	Mean	25.81	26.07
		N	42	42
		Std. Deviation	4.169	4.111
	Total	Mean	25.38	25.60
		N	84	84
		Std. Deviation	4.254	4.206
Female	Control	Mean	26.90	26.98
		N	42	42
		Std. Deviation	3.900	3.885
	Treatment	Mean	25.98	26.36
		N	42	42
		Std. Deviation	4.464	4.282
	Total	Mean	26.44	26.67
		N	84	84
		Std. Deviation	4.192	4.076
Total	Control	Mean	25.93	26.05
		N	84	84
		Std. Deviation	4.219	4.177
	Treatment	Mean	25.89	26.21
		N	84	84
		Std. Deviation	4.294	4.174
	Total	Mean	25.91	26.13
		N	168	168
		Std. Deviation	4.244	4.164

Table E2

Means and Standard Deviations for CIS Confidence Scores

Sex	Group		CIS Confidence pre	CIS Confidence post
Male	Control	Mean	28.31	28.36
		N	42	42
		Std. Deviation	4.719	4.710
	Treatment	Mean	27.98	27.95
		N	42	42
		Std. Deviation	4.442	4.294
	Total	Mean	28.14	28.15
		N	84	84
		Std. Deviation	4.558	4.484
Female	Control	Mean	28.05	28.10
		N	42	42
		Std. Deviation	5.065	4.982
	Treatment	Mean	27.90	28.26
		N	42	42
		Std. Deviation	4.853	5.360
	Total	Mean	27.98	28.18
		N	84	84
		Std. Deviation	4.931	5.144
Total	Control	Mean	28.18	28.23
		N	84	84
		Std. Deviation	4.867	4.821
	Treatment	Mean	27.94	28.11
		N	84	84
		Std. Deviation	4.624	4.830
	Total	Mean	28.06	28.17
		N	168	168
		Std. Deviation	4.735	4.811

Table E3

Means and Standard Deviations for CIS Satisfaction Scores

Sex	Group		CIS Satisfaction pre	CIS Satisfaction post
Male	Control	Mean	27.79	28.17
		N	42	42
		Std. Deviation	3.000	3.428
	Treatment	Mean	27.83	27.64
		N	42	42
		Std. Deviation	3.378	3.341
	Total	Mean	27.81	27.90
		N	84	84
		Std. Deviation	3.175	3.375
Female	Control	Mean	27.93	27.67
		N	42	42
		Std. Deviation	3.564	3.497
	Treatment	Mean	27.45	27.74
		N	42	42
		Std. Deviation	3.451	3.155
	Total	Mean	27.69	27.70
		N	84	84
		Std. Deviation	3.495	3.310
Total	Control	Mean	27.86	27.92
		N	84	84
		Std. Deviation	3.275	3.451
	Treatment	Mean	27.64	27.69
		N	84	84
		Std. Deviation	3.400	3.230
	Total	Mean	27.75	27.80
		N	168	168
		Std. Deviation	3.330	3.334

Appendix F

Means and Standard Deviations for Confidence and Attitude Toward Success Scores on the FSMAS

Table F1

Means and Standard Deviations for FSMAS Confidence Scores

Sex	Group		FS Confidence pre	FS Confidence post
Boy	Control	Mean	37.86	37.50
		N	42	42
		Std. Deviation	8.209	8.417
	Treatment	Mean	37.50	37.45
		N	42	42
		Std. Deviation	8.220	8.079
	Total	Mean	37.68	37.48
		N	84	84
		Std. Deviation	8.167	8.200
Girl	Control	Mean	36.60	36.60
		N	42	42
		Std. Deviation	7.862	7.831
	Treatment	Mean	37.67	37.83
		N	42	42
		Std. Deviation	8.366	7.957
	Total	Mean	37.13	37.21
		N	84	84
		Std. Deviation	8.087	7.871
Total	Control	Mean	37.23	37.05
		N	84	84
		Std. Deviation	8.014	8.093
	Treatment	Mean	37.58	37.64
		N	84	84
		Std. Deviation	8.244	7.972
	Total	Mean	37.40	37.35
		N	168	168
		Std. Deviation	8.107	8.014

Table F2

Means and Standard Deviations for FSMAS Attitude Toward Success Scores

Sex	Group		FS Attitude Toward Success pre	FS Attitude Toward Success post
		Mean	52.52	52.57
	Control	N	42	42
		Std. Deviation	3.373	3.132
		Mean	51.95	51.76
Boy	Treatment	N	42	42
		Std. Deviation	3.092	2.994
		Mean	52.24	52.17
	Total	N	84	84
		Std. Deviation	3.229	3.073
		Mean	52.19	52.31
	Control	N	42	42
		Std. Deviation	3.187	3.294
		Mean	52.24	52.45
Girl	Treatment	N	42	42
		Std. Deviation	2.853	2.787
		Mean	52.21	52.38
	Total	N	84	84
		Std. Deviation	3.006	3.034
		Mean	52.36	52.44
	Control	N	84	84
		Std. Deviation	3.266	3.198
		Mean	52.10	52.11
Total	Treatment	N	84	84
		Std. Deviation	2.960	2.896
		Mean	52.23	52.27
	Total	N	168	168
		Std. Deviation	3.110	3.046

Appendix G

Copyright permission for Figures 1-3

December 19, 2012

Betsy Peisach
Managing Director Education Marketing and Outreach
Maryland Public Television (MPT)
11767 Owings Mills Blvd.
Owings Mills, MD 21117

Dear Ms. Peisach:

I am completing a doctoral dissertation at Lehigh University entitled "The Effects of Digital Games on Middle School Students' Mathematical Achievement." I would like MPT permission to include screenshots from the game "Lure of the Labyrinth" in my dissertation.

The screenshots to be reproduced are:

Figure X. Screenshot from *Lure of the Labyrinth* showing the graphic novel style used throughout the game. Retrieved from http://labyrinth.thinkport.org/www/


Figure Y. Employee cafeteria puzzle screenshot from *Lure of the Labyrinth*. The puzzle shown in this figure furthers the game narrative and requires proportional thinking to solve. Retrieved from http://labyrinth.thinkport.org/www/

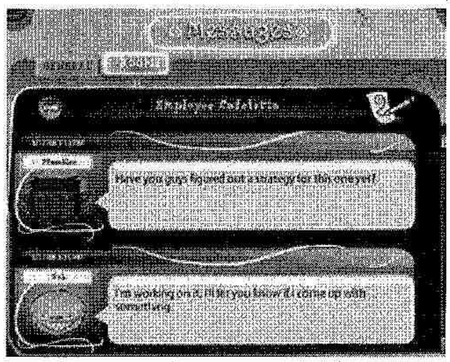

Figure Z. Tasti Pet Communicator (TPC) screenshot from *Lure of the Labyrinth.* The TPC allows players to collaborate electronically while in gameplay mode. Retrieved from http://labyrinth.thinkport.org/www/

The requested permission extends to any future revisions and editions of my dissertation, including non-exclusive world rights in all languages, and to the prospective publication of my dissertation by ProQuest through its UMI® Dissertation Publishing business. ProQuest may produce and sell copies of my dissertation on demand and may make my dissertation available for free internet download at my request. These rights will in no way restrict republication of the material in any other form by you or by others authorized by you. Your signing of this letter will also confirm that MPT owns the copyright to the above-described material and you have the right to provide permission for use.

If these arrangements meet with your approval, please sign this letter where indicated below, scan, and return to me as an email attachment. Thank you very much.

Sincerely,

Pilar Starkey

PERMISSION GRANTED FOR THE USE REQUESTED ABOVE:

Betsy Peisach

Date: 2-22-13

PILAR L. STARKEY

Email: pilarosa@aol.com Cell # 0735601765

SCHOOL ADMINISTRATION

Over 16 years of teaching and administrative experience demonstrating a consistent track record of outstanding performance. Strong qualifications in multiple areas of school leadership and teaching: certified educator, guidance counselor, college placement officer, special needs coordinator and principal. Effective communicator, leader and problem solver who builds teamwork and possesses the ability to translate theory into practice.

RELEVANT EXPERIENCE

Leadership—Rapid advancement through progressively responsible positions in school administration. Key member in multiple education settings in developing new programs, supporting staff and students, fostering increased enrollment and upholding the program's mission.

- Active role in the ongoing accreditation process for a K-12 school
- Instrumental in the over 100% increase in student enrollment for a language school
- Provided support for mission building and sound instructional practice at the collegiate, secondary and primary levels
- Facilitated positive communication amongst stakeholders
- Fostered a positive multicultural environment
- Effectively led community through the initial phases of adopting the IBO programmes
- PA Principal and Superintendent certification pending

Human Resources and Training—Involved in the development, implementation and review of organizational policies and procedures in multiple settings. Supported a learning community atmosphere while providing direction and support to staff and faculty. Designed and implement educational trainings.

- Experienced in delivering training, developing programs, scheduling, evaluating performance and supervision
- Effectively worked with stakeholders in developing, maintaining, and evaluating management practices that support inclusion, excellence, and personal responsibility
- Conducted training seminars both in house, and within the larger community (Peace Corps, US Embassy, AISA, and US school districts)

Teaching and Curriculum Development—Taught a variety of subjects including ESL, social studies, science, art and psychology to students of different ages and abilities. Instrumental in curriculum design, review and implementation.

- Wrote curricula at the K-12, collegiate, and adult education level
- Developed and implemented college readiness programs including PSAT, SAT, TOEFL, and AP preparation
- Actively involved in curriculum mapping and facilitating improvement regarding scope, sequencing, and integration
- K-8 teaching certification (NJ), with experience in Science, ESL and Special needs

CAREER HISTORY

INTERNATIONAL SCHOOL OF KENYA, Nairobi, Kenya (2009-Present)
Teacher – grade 8 humanities, grade 8 science, integrated math, Algebra
Head of Department
AMERICAN SCHOOL OF ANTANANARIVO, Madagascar (2002-2009)
Principal K-12 (2005-2009)
Special Needs Coordinator (2003-2005)
Teacher—Primary to HS Science, MS History, HS Psychology (2002-2009)
Guidance and College Placement Counselor (2002-2009)
ASSOCIATION OF INTERNATIONAL SCHOOLS IN AFRICA (2008-2009)
Virtual School Project Representative for the Africa Region
PSYCHOLOGIST, Vermont, Gabon, Madagascar (1997-2009)
Counseling
Psycho-educational Evaluations
AMERICAN INSTITUTE FOR MASTERING ENGLISH, Gabon (1999-2001)
Curriculum Coordinator (2000-2001)
Teacher (1999-2001)
GREEN MOUNTAIN COLLEGE, Poultney, Vermont (1995-1996 & 1997-1999)
Assistant Director Counseling Services (1998 to 1999)
Counselor and Member of Student Affairs Department (1997 -1999)
Educator (1997-1999)
Intern Counselor and Member of Student Affairs Department (1995-1996)
FARM & WILDERNESS, Plymouth, Vermont (1998 & 1999)
Mental Health Coordinator
Trainer/Educator
ACTING OUT, Keene, New Hampshire (1996-1997)
Community Educator
SPRING LAKE RANCH, Cuttingsville, Vermont (1993-1995, 1996)
Resident Coordinator (1996)
Aftercare Manager (1994-1995)
Residential Advisor (1993-1994)
STARLING ART, Stroudsburg, Pennsylvania (1991-1993)
Owner & Manager
POSTGRADUATE CENTER, New York, New York (1988-1991)
Vocation Counselor
Trainer/Educator

EDUCATION

Doctoral Candidate in Education Leadership • Lehigh University, PA • Present
Post Graduate Certification in International Education • The College of New Jersey • 2005
Masters in Counseling Psychology • Antioch New England Graduate School, NH • 1997
Bachelor of Arts in Psychology • Williams College, MA, 1988

ADDITIONAL SKILLS

Computer literate with System Administrator training in Black Board, MOODLE, Atlas Rubicon, EdAdmin, and Rediker Administrator's Plus. Advanced French. 10+ years living and working in cross cultural situations. Certified scuba diver and artist.