2014

Is it all in the game? Flow experience and scientific practices during an INPLACE mobile game

Denise Marie Bressler
Lehigh University

IS IT ALL IN THE GAME?

FLOW EXPERIENCE AND SCIENTIFIC PRACTICES

DURING AN INPLACE MOBILE GAME

by

Denise M. Bressler

A Dissertation

presented to the Graduate and Research Committee

of Lehigh University

in Candidacy for the Degree of

Doctor of Philosophy

in

Learning Sciences and Technology

Lehigh University

April 2014

CERTIFICATE OF APPROVAL

The dissertation of Denise M. Bressler is approved and recommended for acceptance as a dissertation in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

DATE: ___________________________________

COMMITTEE:

____

Alec M. Bodzin, Ph.D., Chair
Associate Professor/Teaching, Learning, & Technology

____

Scott Garrigan, Ed.D.
Professor of Practice/Teaching, Learning, & Technology

____

Thomas C. Hammond, Ph.D.
Assistant Professor/Teaching, Learning, & Technology

____

Kurt D. Squire, Ph.D.
Full Professor/Educational Comm. & Technology

* Signatures are on file at Lehigh University in the College of Education.

Dedication

I want to dedicate this to my husband…

who is the most incredible man I know and my very best friend

and

to my grandpa…

who was always my biggest fan.

Acknowledgements

I know they say it takes a village to raise children; however, there is a similar truth when it comes to preparing a dissertation. Many people have served a role in this study; some were starring roles, some were supporting roles, but all were significant.

First, I would like to thank my committee. I would like to thank Dr. Alec Bodzin, my advisor and chair, for helping me grow into a scholar; it's been a wonderful journey. I am tremendously appreciative of Dr. Thomas Hammond's early and continued support of my intellectual interest in using mobile augmented reality (AR) learning environments. Dr. Scott Garrigan has been enormously helpful by keeping me on top of the latest in game-based learning as well as tools for creating AR learning experiences. Also, I am thoroughly grateful to Dr. Kurt Squire who created a body of work that inspired my interest in doing this type of research.

Second, I would like to thank the administrators and the teachers that participated in this project. The administrators believed in this project and supported it wholeheartedly—working in their district was a dream-come-true for an aspiring educational researcher. The teachers got excited about disturbing the inertia of their classrooms and gave me full reign to do so; they also gave up endless personal prep periods and made me feel completely at home in their classrooms.

Third, I would like to thank my family. I am forever indebted to my husband, Zach, for his unconditional love, unwavering support, sincere interest in my work, and for also doing the laundry.[1] He made this dream a possibility by believing in my passion and participating in an egalitarian marriage—even if that's (not yet) the norm. I want to thank my daughter, Ayla, who talked with mommy about statistics, discussed my project pictures, and came up with a very useful metaphor about honey and vinegar to encapsulate the importance of productive discourse

[1] when professor Rosabeth Moss Kanter of the Harvard Business school was asked what a man could do to best support women's leadership – she said – "the laundry" (Sandberg, 2013).

in collaborative learning; she is so much more insightful about the human condition than anyone her age.[2] My son, TJ, was not even born when I started this degree program, and may not remember much of mommy's dissertation, but he was always there to come squish my face when I needed a little love pick-me-up; he always kept me smiling. Also, I cannot thank my mom and Bill enough for supporting me not only in my schoolwork but also in my life; they looked after my kids on many occasions making it possible to complete my schoolwork. My mom was especially heroic because she listened, and Listened, and LISTENED to my endless brain dumps and mindless ramblings without ever making me feel guilty. My dad has had a great impact on my work as well by introducing me to the writings of his educational heroes, such as Seymour Papert, and by seamlessly incorporating computers into my childhood; thank you for constantly supplying me with references, links, and interesting TED talks to keep me thinking, learning, and making connections to the bigger picture of educational reform. Additionally, my in-laws have also helped me along this exciting yet arduous road…namely Jill, Sandy, and Peter.

Finally, I owe a huge debt of gratitude to my colleagues. Bill Farina was assigned as my research assistant a few weeks before data collection and enthusiastically jumped into the fire; I would not have survived without him. To the Lovely Ladies of Lehigh: you are my cheerleaders, confidantes, and most enthusiastic supporters. One specific shout-out to Farah for being the best study buddy a girl could have. I have also enjoyed a wealth of encouragement from the ARIS community…so thank you to a very cool group of super smart folks. To my statistical superhero…thanks for being a great friend through all of this. Ultimately, to all the unnamed associates who helped me along the road to this achievement, thank you for joining me on this incredible adventure…I love you all.

[2] Ayla is 7.5 years old; she is in second grade.

TABLE OF CONTENTS

List of Tables

List of Figures

ABSTRACT

Mobile science learning games show promise for promoting scientific practices and high engagement. Researchers have quantified this engagement according to flow theory. Using an embedded mixed methods design, this study investigated whether an INPLACE mobile game promotes flow experience, scientific practices, and effective team collaboration. Students playing the game (n=59) were compared with students in a business-as-usual control activity (n=120). Using an open-ended instrument designed to measure scientific practices and a self-report flow survey, this study empirically assessed flow and learner's scientific practices. The game players had significantly higher levels of flow and scientific practices. Using a multiple case study approach, collaboration among game teams (n=3 teams) were qualitatively compared with control teams (n=3 teams). Game teams revealed not only higher levels of scientific practices but also higher levels of engaged responses and communal language. Control teams revealed lower levels of scientific practice along with higher levels of rejecting responses and command language. Implications for these findings are discussed.

CHAPTER 1: INTRODUCTION

Education in the 21st Century

For the United States to compete in the global economy it must optimize its knowledge-based resources; this is particularly true in the fields of science and technology (National Research Council, 2007a). To bolster the resources feeding into the knowledge economy, the nation needs strong schools with high graduation rates. Unfortunately, too many students drop out of high school, leaving the graduation rate unacceptably low (National Research Council, 2011b). Dropping out is a particular problem among students in low socio-economic areas (Finn, 1989). To understand the dropout problem, the National Research Council (2011b) reviewed a growing body of student engagement studies and concluded that “dropping out is but the final stage in a dynamic and cumulative process of disengagement from school” (p. 61).

Disengagement is understandable in a culture of schooling that is entrenched in didactic teaching styles and the drudgery of boring seatwork (Gatto, 2009; Rathunde & Csikszentmihalyi, 2005). Over 20 years ago, a report called *A Nation at Risk: The Imperative for Educational Reform* warned that students were not being challenged to use higher-order thinking skills (National Commission on Excellence in Education, 1983). Not much has changed. With the No Child Left Behind Act of 2001, schools have shifted their curricula to emphasize more rote memorization as a means to prepare for standardized tests (Ellis, 2007). This has reduced the amount of meaningful learning taking place in schools (Kohn, 2000; Ravitch, 2011). Now the nation is at greater risk.

Not only are students steadily disengaging as they move through the educational system, but those who do graduate are not fully equipped for careers in the 21st century. The Partnership for 21st Century Skills (2013) has defined the skills, knowledge, and expertise necessary for today's work-life; critical thinking, problem solving, communication, and collaboration are of critical importance. These skills are not cultivated in the majority of U.S. schools, yet they are essential to garnering success as adults (National Research Council, 2012b) in today's "knowledge economy" (Powell & Snellman, 2004).

According to a report commissioned by the National Research Council (2012a), hereinafter referred to as NRC, the ability to collaboratively solve problems is of the utmost importance in scientific careers. According to K-12 science framework authored by the NRC (2012a), "science is fundamentally a social enterprise, and scientific knowledge advances through collaboration and in the context of a social system with well-developed norms" (p. 27). The job prospects in science and technology are growing (Lockard & Wolf, 2012); however, our students are underprepared for the job requirements. When we do not prepare our students adequately for the workplace, our national prosperity suffers (NRC, 2012b). To make the US globally competitive in science and technology, students need to be engaged with science education, engage in authentic scientific practices, and learn to collaborate successfully.

Promoting Scientific Practices and Collaboration with Mobile Games

For the United States to continue to compete in the global economy, our schools need to provide learning environments that engage students, promote scientific practices, and encourage effective collaboration. Well-designed educational games hold great promise to satisfy this extraordinarily unique but critically important set of requirements. James Paul Gee, a preeminent scholar on games and learning, argued that while school has set aside some of the best learning principles, games have embraced and incorporated them (Gee, 2007).

Hard Fun in Mobile Games Promotes Engagement

Games promote engagement through hard fun. Games embrace the idea of 'hard fun' as discussed by Seymour Papert (1996), a concept first posited by Dewey (1910/1997) who stated that "to be playful and serious at the same time is possible, and it defines the ideal mental condition" (p. 218). Hard fun means exactly that—the activity is both *hard* and *fun* simultaneously. In hard fun, the fun offers emotional benefits. According to game designer Jane McGonigal (2011), gamers "would rather work hard than be entertained" because it contributes to a positive mood (p.32).

When people experience hard fun, they are engaged. According to Annetta and Cheng (2008), "educational games seem to be effective in enhancing motivation and increasing student interest in subject matter…[they] generate a much higher level of students' positive emotional engagement" (p. 2). High levels of engagement have been documented in early research on collaborative, mobile augmented reality learning games (Dunleavy, Dede, & Mitchell, 2009; Facer et al., 2004; Perry et al., 2008; Schwabe & Göth, 2005). More recently, researchers have found that such engagement is related to flow (Admiraal, Huizenga, Akkerman, & Ten Dam, 2011; Bressler & Bodzin, 2013; Park, Parsons, & Ryu, 2010; Ryu & Parsons, 2012). Flow is a positive psychological state that is challenging, intrinsically rewarding, and enjoyable (Csikszentmihalyi, 1975). Specifically, collaborative mobile games hold promise for engaging students by promoting flow during gameplay.

Situated Cognition in Mobile Games Promotes Scientific Practices

Games promote scientific practice through situated cognition, a theory that is considered a "work in progress" (Kirshner & Whitson, 1997), but has a history dating back to Dewey, Montessori, Piaget, and Vygotsky. In 1989, Brown, Collins, and Duguid published the first article on situated cognition asserting that learners cannot separate what they learn from how they learn it. Games afford students this type of active participation in their learning—players *learn by doing*. Additionally, the distinction of situated cognition is that the knowledge to be learned is deeply rooted within the social process (Lave & Wenger, 1991). According to the report commissioned by the NRC (2012a), science is a "social enterprise," so for science learning, collaborative games can provide context and tools in which to conduct scientific practice; players can learn science by doing science.

In 2011, the NRC published the report *Learning Science through Computer Games and Simulations* and highlighted that certain "games have potential to advance multiple science learning goals" (p. 2). Collaborative mobile learning games, still in infancy, show potential for promoting science learning (Squire & Klopfer, 2007) and scientific literacy (Squire & Jan, 2007). Mobile games designed to put students in professional roles have shown that during gameplay students engage in the process of scientific inquiry (Dunleavy et al., 2009; Rosenbaum, Klopfer, & Perry, 2007; Squire & Klopfer, 2007) and argumentation (Mathews, Holden, Jan, & Martin, 2008; Squire & Jan, 2007). Squire and Jan (2007) stated that these "games could have promise for tools that develop scientific literacy" (p. 23). Specifically, collaborative mobile games hold promise for promoting scientific practices during gameplay.

Social Constructivism in Mobile Games Promotes Collaboration

Games can promote social constructivism by promoting collaboration, specifically collaborative discourse. Well-designed collaborative games support social constructivism, a theory mainly associated with Vygotsky although it draws influences from Piaget. As summarized by Mooney (2000), Vygotsky connected children's play to their social learning; as children listen and speak to each other, their social interaction directly contributes to their knowledge construction, and they learn from each other. Piaget (1985) posited that such exchanges present opportunities for intellectual growth as children experience disequilibrium between their own beliefs and those of their classmates.

Research on collaborative educational games has shown that gameplay positively impacts the development of collaboration skills (Sánchez & Olivares, 2011) and player's perceptions of their social interactions (Mansour & El-Said, 2009). Specifically, students enjoy playing collaboratively because it encourages discussion among players (Sharritt, 2008). The sociocultural learning that takes place within the game works best when there is shared power and authority through scripted collaboration (Demetriadis, Tsiatsos, & Karakostas, 2012). Driscoll (2005) summarized Vygotsky's requirement that social interaction must have intersubjectivity to promote effective learning, explaining that students must "co-construct the solution to a problem or share in joint decision making" (p. 258). Several researchers in the field of mobile learning games incorporated this concept into their game designs by creating interdependent roles (Bressler & Bodzin, 2013; Dunleavy, et al., 2009; Dunleavy & Simmons, 2011; Klopfer, Perry, Squire, & Jan, 2005). As summarized by Klopfer (2008), students playing collaborative mobile learning games "help each other, observe each other, and act together to create communities as they learn to solve problems" (p. 223). Overall, research indicates that collaborative mobile games hold promise for promoting effective collaborative practice by scaffolding and supporting discourse during gameplay.

Statement of Purpose

The main purpose of this research study is to expand and extend the knowledge documented by the mobile learning games community in education. In order to make a contribution to the knowledge base, this study will use a formal context, with empirical methods, and a larger sample size.

- First, previous research on mobile learning games has utilized informal contexts such as museums (Cabrera et al., 2005; Klopfer, Perry, Squire, Jan, & Steinkuehler, 2005), zoos (Perry et al., 2008), nature centers (Klopfer & Squire, 2008), and libraries (Czarnecki, 2012). Studies that have taken place in the formal school setting either used non-content classes like physical education (Spikol & Milrad, 2008) or worked with students outside of normal class time (Bressler & Bodzin, 2013). Findings indicate promise in formal educational settings, yet few have empirically studied mobile games in the formal learning environment. Therefore, this study will take place during regular class meeting times, within the school building. Furthermore, the activities that are part of the study are aligned to the adopted school curriculum.
- Second, most research teams have used some form of design-based research (Barab, 2006). This type of research approach has enabled the field to explore practical issues pertaining to using mobile AR games in the K-12 environment as well as what design elements are most successful (Dunleavy & Dede, 2013). The field has learned that playing professional roles supports scientific practices (Dunleavy et al., 2009; Mathews et al., 2008; Rosenbaum et al., 2007; Squire & Jan, 2007; Squire & Klopfer, 2007), differentiated, interdependent roles effectively scaffold collaborative problem solving (Dunleavy et al., 2009; Lai & Wu, 2006, Squire & Jan, 2007), and narrative-based games provide a structure in which to think (Squire & Klopfer, 2007). Now that the field has established a baseline understanding of best practices for mobile AR games, the next logical step is to develop an empirical study that employs such design principles. As will be discussed in the next chapter, the best practices of

collaborative mobile games have generated a need for a new category of Serious Educational Games: INPLACE mobile games. The author of this dissertation created the INPLACE acronym and it stands for: **I**nterdependent, **N**etworked, **P**articipatory **L**earning, **A**ugmented, **C**ollaborative **E**xperience. For this study, an INPLACE mobile game will be created and then implemented with a quasi-experimental group; the study will also examine a control group to understand the impact of the intervention.

- Third, many previous research studies have worked with small sample sizes, sometimes only 10 or 20 students (Facer et al., 2004; Klopfer, Sheldon, Perry, & Chen, 2012). In order to generalize the findings to a larger population, this study will have a larger sample size of approximately 300 students.

The secondary purpose of this research study is to empirically test whether a flow-like state experienced within a mobile game leads to greater learning gains and if those learning gains are consistent for all levels of students. Researchers with extensive experience in game-based learning have all commented on the connection between games and flow theory as defined by Csikszentmihalyi (1990). People often feel flow when they are involved in an activity that is challenging, intrinsically rewarding, and enjoyable (Csikszentmihalyi, 1975). Eric Klopfer (2008) stated that the same qualities that are in activities that lead to flow are found in engaging games. Kurt Squire (2011) stated that the sequencing of compelling games helps to induce flow in players. Katie Salen and Eric Zimmerman (2004) specified that the components of flow as outlined by Csikszentmihalyi have clear parallels with games and that "games are one of the best kinds of activities to produce flow" (p. 338). The experts believe that there is a connection between flow and games, but does the experience of flow lead to greater learning?

The work of Csikszentmihalyi (1990) indicated that flow activities do lead to higher performance. Specifically, Csikszentmihalyi (1990) stated that "flow activities lead to growth and discovery" (p. 75). Growth occurs because proficiency at a skill leads to boredom, and boredom is contradictory to flow. Thereby, to stay in flow, participants need to continue growing and discovering new challenges. Some researchers have found that learners who experience flow have improved attitudes towards learning (Faiola, Newlon, Pfaff, & Smyslova, 2013), while others support the notion that playing mobile learning games could increase learning (Klopfer et al., 2012). No prior research has shown empirical data to prove that flow increases learning.

The tertiary purpose of this research study is to take a closer look at socially constructed knowledge that takes place within the collaborative discourse of players. Research shows that effective group communication is a key predictor of group success (Barron, 2003; Pentland, 2012). Therefore, this study will investigate if there are successful communication patterns in the responses between players. Furthermore, as outlined earlier, there is reason to believe that a well-designed mobile game can support scientific practice through scaffolded collaborative discourse. By investigating how scientific practice unfolds during player discourse, specific data will be gathered to further understand this notion.

Overall, given the paucity of research on implementing mobile learning games in formal classrooms and the promising nature of mobile games to foster scientific practices and effective collaboration, the purpose of this research study is to support the idea that a classroom science investigation framed within an INPLACE mobile game provides opportunities for middle school students to learn scientific practices and collaborate more successfully. This study will investigate not only the learning and collaboration of those playing the game but also of those participating in a similar non-game-based investigation in the classroom. Student learning as a result of the experiment and control will be assessed along with flow experience and collaborative discourse.

This study will employ a mixed methods approach with an embedded design. According to Creswell and Plano Clark (2011), mixed methods research is “an intuitive way of doing research” (p. 1). To tell the complete story, a qualitative strand will be embedded within the quantitative experiment. Qualitative analysis is necessary to interpret the quantitative results; furthermore, while science practices and flow experience can be empirically assessed, collaborative discourse requires a qualitative approach.

Research Questions

This study will compare a control activity with a quasi-experimental activity on two dependent variables including flow experience and a culminating artifact that measures scientific practices. First, the study will investigate:

1. Do students in the INPLACE mobile game compared to the control activity report a more substantive flow-like experience and create culminating artifacts with more evidence of scientific practices?
    a. Are there any differences between treatment groups for flow experience or culminating artifact score when investigated separately?
    b. Do reading achievement scores influence any of the above relationships?
2. Does flow experience predict culminating artifact score?

Second, the study will qualitatively assess collaborative discourse during both the quasi-experimental activity and the control activity.

3. How do in-team communication responses during the INPLACE mobile game compare to in-team responses during the control activity?

4. How do scientific practices of teams in the INPLACE mobile game compare to those of teams in the control activity?
5. How else are treatment groups different when qualitatively analyzed at the team level?

Significance of this Study

The 2011 NRC report on games and simulations explained that research on student learning with science games is promising but inconclusive; it called for more research to develop scalable designs and to increase understanding of the role games can play in science learning. The goals of this study are to 1) contribute new pedagogical and game design knowledge to increase the field's understanding of mobile games and science learning and 2) create a scalable design that educators can easily replicate and implement. Therefore, the proposed study will make a contribution both to theory and to practice.

This study will do more than increase the knowledge base concerning mobile games and learning—this study aims to find a way to engage our students in science learning. The nation needs to deeply engage students in their schooling, particularly in science and technology. Improving engagement is important for decreasing drop-out rates, but more importantly, it's a prerequisite for productive learning. As stated by Csikszentmihalyi, Rathunde, and Whalen (1993), "learning has to be engaging and rewarding for students to learn" (p. 9).

Finally, this study not only aims to create a scalable model but also a model that effectively promotes crucial 21st century skills such as collaboration and communication. Creating a scalable game design is an important need of the educational community since experts foresee K-12 education incorporating game-based learning within two to three years (Johnson, Adams, & Cummins, 2012). However, creating a scalable model that effectively promotes 21st century skills means that, if the model is widely used, our students can become fully prepared to contribute to today's knowledge-based workforce.

CHAPTER 2: REVIEW OF THE LITERATURE

Saving American Schools from Obsolescence

In the past century, we have seen a societal paradigm shift that has rendered the traditional educational system obsolete. In the Industrial Age, leadership was autocratic and factory workers were compliant (Reigeluth, 1994). The factory model of education was sufficient at providing students with enough knowledge and skills to live out full lives as factory workers. However, with the onset of the Information Age, the workplace is evolving (Kanter, 2001). Mass collaboration is changing how business is conducted: leadership is shared; competitive principles include openness and sharing; and knowledge workers expect a participatory democracy (Tapscott & Williams, 2006). As discussed by Thomas and Brown (2011), the world is changing—and in accordance—so is knowledge. We need a system that prepares students for working in the Information Age, not the Industrial Age (Reigeluth, 1994).

The Partnership for 21st Century Skills (2013) has defined the skills, knowledge, and expertise necessary for today's work-life; critical thinking, problem solving, communication, and collaboration are essential. Research shows that the workforce is drastically ill-prepared in such expertise; companies need to train their workers in not only 21st century skills but also basic skills like reading and math (Casner-Lotto, Rosenblum, & Wright, 2009). For the United States to compete in the global economy, it must optimize its knowledge-based resources; Diane Ravitch (2011), a historian of education, succinctly stated that "education is the key to developing human capital" (p. 223). With our human capital so ill-prepared, we need to hold up a magnifying glass to the educational system.

Cultivating human resources is particularly urgent in the fields of science and technology (NRC, 2007a). Low level skill sets will not solve critical scientific issues such as climate change and the energy crisis; the country's most pressing problems need innovative problem solvers

(Gee, 2010). According to the report published by the NRC (2012a), “science is fundamentally a social enterprise, and scientific knowledge advances through collaboration and in the context of a social system with well-developed norms” (p. 27). The ability to collaboratively solve problems is of the utmost importance in scientific careers. Unfortunately, students are not prepared for such collaborative work because school environments place too much emphasis on individual work (Bransford, Brown, & Cocking, 2000) and standardized testing (Ravitch, 2011).

Beyond lack of effective preparation, kids are bored at school (Gatto, 2009; Willingham, 2010), and those who become completely disengaged drop out (NRC, 2011b). Young students learn differently—they want active, exploratory learning (Kirriemuir & McFarlane, 2004). Due to today’s culture of high-stakes testing, most schools still place emphasis on content mastery; students are acquiring isolated skills and specific facts rather than spending time on more meaningful, learner-driven processes such as scientific practices and conceptual thinking (Kohn, 2000; Ravitch, 2011; Richardson, 2012). Don Tapscott (2009), an expert on the transformative power of the digital age, posits that young students who are growing up digitally connected learn differently; because their brains have developed differently, their brains process information differently. This *different* style of learning is largely why students are so bored at school (Goyal, 2012); content mastery needs to be replaced by learning mastery (Richardson, 2012). Additionally, students are surrounded by technologies outside of school (Ito et al., 2010; Lenhart, Hitlin, & Madden, 2005)—technologies that support their neo-millennial learning styles (Dieterle, Dede, & Schrier, 2007). Curricula struggle to support tech-savvy students (Arafeh, Levin, Rainie, & Lenhart, 2002) because a tension exists between traditional, didactic teaching approaches and the distributed, decentralized nature of today’s media (Roschelle & Pea, 2002).

Education has always struggled to integrate new technologies into the classroom. For decades, progressive educational reformers have held up technology as the way to save our schools. Classrooms saw mainframes and minicomputers in the 1960s, personal computers in the 1970s and 1980s, and multimedia technologies and eventually the internet in the 1990s (Wenglinsky, 2005). Technology reform movements have not historically worked though, because having the technology in the classroom and using it effectively are two different notions. In 2001, Larry Cuban studied classroom use of new technologies and determined that putting computers and internet access in classrooms did not dramatically change education (Cuban, 2001). Several years later, the Department of Education (DOE) conducted a study to determine whether student scores in math and reading increased when teachers had access to software designed to improve reading and math; their findings were not positive either (Dynarski et al., 2007). On the other hand, after eight years of data collected from Apple's longitudinal study of the Apple Classrooms of Tomorrow (ACOT), not only did scores demonstrate a marked increase but students were exhibiting skills in inquiry, collaboration, and problem solving well beyond students in traditional classrooms (Dwyer, 1994). The ACOT showed significant differences because they let the technology "disturb the inertia" of the traditional classroom. Norris and Soloway (2011) recently reminded the educational community that "as long as schools use computers to teach the existing curriculum using existing pedagogies little to nothing will be gained" (p. 4) as was observed by Cuban and the DOE. Apple saw results because Apple did things *differently*. First, students actually used the new technology. Second, and more importantly, the teachers and technologists did not try to do old things with a new tool, they let the new tool dictate new possibilities; therefore, students' technology use enhanced their learning.

For new technologies to be successful in the classroom, educators needs to break out of the old model of thinking and reframe the mindset; you cannot restructure a “horse-and-buggy” system into a “spacecraft” model (Banathy, 1994, p.27). According to Davidson and Goldberg (2010), “new technologies are changing how people of all ages learn, play, socialize, exercise judgment, and engage in civic life,” yet learning institutions have barely changed in accordance (p. 12). Apple approached ACOT with a new mindset and had students use computers in the classroom in ways uniquely afforded by the computer. In order to bring the educational system out of obsolescence, Davidson and Goldberg (2010) argued that formal learning institutions need to take advantage of new modes of participatory learning enhanced by digital media.

The new digital medium that warrants the most educational potential at this moment is mobile technology, because we are at a tipping point (Gladwell, 2000). For years, mobile phones have been a vital part of teenager’s social lives (Katz, 2006); young people live in a mobile world and their devices connect them to it. Mobile technology has had time to incubate outside the K-12 system; now, it is ready to be a “disruptive innovation” (Christensen, Horn, & Johnson, 2008). More and more schools are inviting students to *bring your own device* (Norris & Soloway, 2011). We need to create and enact curriculum activities that ensure students will garner important skills and benefit from the unique affordances of mobile technology—no more business-as-usual.

The body of research is growing that demonstrates how mobile technology can promote the type of skills desperately needed by those entering the workforce. Research supports the effectiveness of mobile activities to scaffold collaborative problem-solving (Bressler & Bodzin, 2013; Dunleavy et al., 2009; Lai & Wu, 2006; Squire & Jan, 2007). Mobile devices also afford portability making them effective tools for situated learning (Dede, 2009; Fotouhi-Ghazvini, Earnshaw, Robison, & Excell, 2009; Rosenbaum et al., 2007) and science inquiry (Chang, Chen, & Hsu, 2011; Looi et al., 2011). Integrating mobile technologies into the classroom holds promise, but to be successful, it requires rethinking "business as usual" modes on instruction because mobile technologies support collaborative learning and active, experiential learning.

Game-based learning is one way to capitalize on the affordances of mobile technology while also resolving some of the issues that have led to the educational system's obsolescence. First, game-based learning is engaging, and "schools need to create intrinsically engaging methods for learning" (Christensen et al., 2008, p. 7). Everyone has a drive to learn; it starts in infancy (Gopnick, Meltzoff, & Kuhl, 1999), yet today's schools are boring to students. The K-12 educational system needs to figure out how to help students discover that drive. According to Annetta and Cheng (2008), "educational games seem to be effective in enhancing motivation and increasing student interest in subject matter…[they] generate a much higher level of students' positive emotional engagement" (p. 2). Second, well-designed collaborative mobile games offer a unique way for students to socially construct knowledge. According to Klopfer (2008), students playing collaborative mobile learning games "help each other, observe each other, and act together to create communities as they learn to solve problems" (p. 223). Third, Serious Educational Games (SEG) for science education are showing promise (NRC, 2011a) and building a body of research towards best practices.

Educational games have long been criticized, but this chapter will make the case that well-designed games make for excellent learning environments, especially for skills that today's workplace demands. Well-designed games also offer an enjoyable and effective way to learn, as well as a way to engage students in their learning. Diane Ravitch (2011), an original proponent of the No Child Left Behind Act of 2001 who has since changed her position, concluded that education needs to offer "plenty of opportunity for children to engage in activities and projects that make learning lively" (p. 13). Games research and educational games are nothing new; in fact, back in 1995 the NRC published the report titled—*Reinventing Schools: The Technology is Now!*—advocating that educators explore the educational potential of technologies widely used outside of school, particularly video games. Today, the widespread use of mobile technology and its inherent affordances make collaborative mobile games part of a viable solution to our current educational crisis.

This chapter will first detail the history and learning theories encapsulated in game-based learning. Next, exemplars of games in K-12 education will be presented along with criticism of educational games. Then, the best practices of game-based learning will be discussed highlighting how games can engage learners, promote scientific practices, and support collaboration through collaborative discourse. Finally, the chapter will conclude with the promise of INPLACE mobile games and posit how such a game can improve middle school science learning.

History and Pedagogy of Game-Based Learning

Games are not a new societal concept; there is evidence that a game similar to today's modern backgammon was played in Mesopotamia during the 26th century BCE (Somervill, 2010). The idea of games as an educational enterprise, although certainly a more recent concept, is not entirely new either (Shelton, 2008). The discussion of games and learning has a long history dating back to Dewey (1910/1997) and Montessori (1912). Even with 100 years of longevity, there is still a lack of cultural acceptance of games as a medium for learning (Kirriemuir & McFarlane, 2004). First, this section will track the historical progression of the idea of games for learning, starting with the early ideas of learning through play. Second, this section will draw the linkages between viewpoints of the theorists and the design of effective game-based learning environments to demonstrate why games for learning are worthy of not only educational attention but also cultural acceptance.

Learning through Play

For those who study play, the idea that people do not innately see the correlation between play and learning is hard to fathom. Stuart Brown (2009), a contemporary expert on play, stated "play isn't the enemy of learning, it's learning's partner. Play is like fertilizer for brain growth. It's crazy not to use it" (p. 101). The business world knows the importance of play as well. Michael Schrage (2000) identified common patterns and practices of the best companies like Disney and IDEO and he concluded: "you can't be a serious innovator unless you are willing and able to play" (p. 2). Ultimately, the idea of learning through play is routed in evolutionary biology. Chris Crawford (1984/2011), one of the pioneers of the computer game design field highlighted this notion in his book explaining that games are an ancient, time-honored, and very effective, educational tool. To prove his point, Crawford used lion cubs as a primary example:

> …cubs are not carefree. They do not indulge in games to while away the years of their cubhood. These games are deadly serious business. They are studying the process of hunting, the skills of survival. They are learning how to approach their prey without being seen, how to pounce, and how to grapple with and dispatch prey without being injured. They are learning by doing, but in a safe way. Better to make mistakes with the butterfly and sibling than with the horns of the wildebeest. (Crawford, 1984, loc.340)

Early ideas of learning through play were published over 100 years ago. In 1910, John Dewey published *How We Think* and discussed the intellectual ideal—to have a balance of playfulness and seriousness. To Dewey (1910), free play in childhood is mental play for the mind; play is not "fooling." In such mental play, thoughts are not careless but rather the course is directed and the process iterative. Montessori addressed this unique paradox of order in play when she discussed games. One of the earliest mentions of games in the context of learning appears in the seminal work of Maria Montessori (1912), *The Montessori Method*:

> "We speak, it is true, of games in education, but it must be made clear that we understand by this term a free activity, ordered to a definite end; not disorderly noise, which distracts the attention." (Montessori, 1912, p.181)

Echoing the idea set forth by Montessori, Johan Huizinga (1938/1971), a Dutch historian, discussed the seriousness of play in his book *Homo Ludens*. Huizinga made the case that one of the most positive features of play is that it creates order, it has rules. He used children's play to exemplify his point, such as when someone deviates from the rules of a game, it spoils the game. Huizinga's book is one of the earliest known academic evaluations of play; he was the first person to truly pontificate on the cultural value of play. Furthermore, Huizinga (1938/1971) hypothesized about whether all science is a game and—while he discarded that thought—he concluded that "the scientist's continued penchant for systems tends in the direction of play" (p. 203).

Recapitulating and extending the ideas of his predecessors, Piaget (1962) posited not only that there was a connection between play and learning but also that learning through play was an extremely important avenue of learning. Piaget believed that a child's insatiable curiosity drives their learning; furthermore, while playing, a child can initiate and control the activity. Ultimately, during play, children assimilate their learning (Piaget, 1962).

Similar to his contemporaries, Vygotsky (1966) also believed that play is not just for pleasure, that learning takes place during children's play. As summarized by Mooney (2000), Vygotsky connected children's play to their social learning; as children listen and speak to each other, their social interaction directly contributes to their knowledge construction, and they learn from each other. In addition to promoting a child's social development, Vygotsky (1978) posited that play is an opportunity for children to develop abstract thought. Overall, Vygotsky (1978) believed that a child in play can achieve great things; he concluded that "play creates a zone of proximal development of the child" (p. 102).

Bruner's (1961) ideas of play support his overarching idea that a learner should become "as autonomous and self-propelled a thinker" as possible (p. 23). To Bruner (1983), when children are in a state of playing, they can be "free and inventive", thus play cultivates a child's creative mind (p.64). Overall, play can be viewed "as a means of improving the intellect" (Bruner, 1983, p.63). Given Bruner's ideas of play, one can see how he would draw this conclusion about the power of games for learning:

> "games go a long way toward getting children involved in understanding language, social organization... they provide a superb means of getting children to participate actively in the process of learning - as players rather than spectators." (Bruner, 1966, p. 95)

Learning Theories that Support Game-Based Learning

As outlined above, developmental psychologists and educational theorists have discussed how children learn through playing for over a century. A leading scholar on games and learning, Kurt Squire (2011) has stated that "in many respects, the promise of video games is about realizing age-old visions of education proposed by Maria Montessori or John Dewey" (p.15). Game-based learning certainly has its skeptics (Clark, 2007; Healy, 1998); however, this section will detail how well-designed games encapsulate extremely effective learning theories.

Researchers in the games and learning field are in agreement that games can support meaningful learning. Though there are several theories that can explain how this complex learning takes place, there are a specific few relevant to this study. First, in a literature review of game based learning, Kirriemuir and McFarlane (2004) determined that a key theme among learning games was the idea of learning by doing, the colloquial term for situated cognition.

Second, Cheng and Tsai (2012) concluded that profitable frameworks for future augmented reality (AR) research in science education included not only situated cognition but also social constructivist learning. Finally, Salen and Zimmerman (2004) discussed the complexities involved in the pleasure often experienced by avid gamers and settled on the theory of flow (Csikszentmihalyi, 1975) as the best explanation: "we believe there is an intrinsic connection between game play and flow" (p. 338). While flow theory is not a learning theory, learners who achieve flow are fully focused and challenged, driven by intrinsic motivation and "students will learn only if they are motivated" (Csikszentmihalyi et al., 1993, p. 195).

Hard fun promotes flow. All the theorists above discussed play and learning, denoting that play was not foolish, but in fact, serious and orderly. Seymour Papert (1996) took those ideas and condensed them down to two words: *hard fun*. According to Papert (2002), everyone enjoys hard, challenging things to do, providing the challenge is well-matched to the individual. Papert's description of hard fun is almost synonymous with the notion of challenge-skill balance outlined by Csikszentmihalyi (1975) in his theory of flow. Flow is a positive psychological state that is challenging, intrinsically rewarding, and enjoyable (Csikszentmihalyi, 1975). While there are several characteristics of flow, "flow tends to occur when a person's skills are fully involved in overcoming a challenge that is just about manageable" (Csikszentmihalyi, 1997, p. 30); in other words, when a person is engaged in hard fun, flow can occur.

Games promote engagement by offering experiences of hard fun that promote flow. Hoffman and Nadelson (2010) summarized the existing research on achieving flow while gaming and concluded that "gaming promotes intrinsic motivation, positive affect, and many aspects of the flow experience" (p. 248). Researchers have started using flow as a measure of student engagement with educational computer games (Grimley, Green, Nilsen, Thompson, &

Tomes, 2011; Inal & Cagiltay, 2007; Kiili & Lainema, 2008) and more recently mobile learning games (Admiraal et al., 2011; Bressler & Bodzin, 2013; Park et al., 2010; Ryu & Parsons, 2012). Bressler and Bodzin (2013) found that players desired higher levels of performance because game characteristics seemed to promote the right challenge-skill balance for players to experience flow. Since players in flow desire greater challenges, there is reason to believe that flow is correlated with learning, but further research is needed to investigate the relationship.

Social constructivism. Social constructivism is a learning theory mainly associated with Vygotsky, although it draws influences from Piaget. Piaget believed that children construct knowledge through personal experience; Vygotsky believed that children begin to understand new concepts as they converse with each other—Vygotsky put importance on children's interactions with each other (Mooney, 2000). Vygotsky theorized that within these interactions, social and cognitive development could work together, build on each other, and promote learning. Although Vygotsky would certainly endorse conversation among learners, talking was not enough; sociocultural learning was an interactive experience requiring shared power and authority (Driscoll, 2005). For a collaboration to be an effective learning interaction, it needed to have intersubjectivity: learners must "co-construct the solution to a problem or share in joint decision making" (Driscoll, 2005, p. 258). Piaget (1985) posited that such exchanges present opportunities for intellectual growth as children experience disequilibrium between their own beliefs and those of their classmates. Growth happens because children seek new directions of thought as they try to restore cognitive equilibrium.

Games can promote social constructivism by promoting collaboration, specifically collaborative discourse. Researchers have found that peer collaboration enhances learning while students play educational computer games (Chatterjee, Mohanty, & Bhattacharya, 2011).

Specifically, J.C. Hertz, author of *Joystick Nation: How Computer Games Ate Our Quarters, Won Our Hearts and Rewired Our Minds,* posited that "the discourse is where the learning happens" (Foreman, 2004, p. 58). Now game researchers are starting to have the data to prove that learners are truly constructing knowledge socially as they engage in discourse during gameplay. Steinkueler (2006) confirmed that players in a virtual game environment participated in a discourse; it was created, maintained, and transformed during gameplay. Squire and Jan (2007) studied the discourse between players in a mobile AR game that was designed with intersubjectivity; they contended that players engaged in scientific argumentation because of the task, roles, embedded resources, context, and encompassing activity system of the game.

Situated cognition. The influence of Piaget and Vygotsky can also be seen in the foundations of situated cognition. Although the theory of situated cognition is considered a “work in progress” (Kirshner & Whitson, 1997), it has history dating back to Dewey, Montessori, Piaget, and Vygotsky. All these theorists believed that education should be active and involve social interactions (Mooney, 2000). For example, Vygotsky (1978) stated that an essential component of learning is that “the child is interacting with people in his environment and in cooperation with his peers” (p. 90). This quotation bears similarity to contemporary thinking about situated cognition. In 1989, Brown et al. published their notion that learners cannot separate what they learn from how they learn it, an approach colloquially referred to as *learning by doing*. The distinction of situated cognition is that social interaction is not just an effective means for knowledge construction—as explained by social constructivism—but that the knowledge to be learned is deeply rooted within the social process. According to Lave and Wenger (1991), situated learning is “configured through the process of becoming a full participant in a sociocultural practice” (p. 29). This theory spawned from the idea of learners as apprentices. Just as an apprentice learns by observing, following, and repeating the practices of her mentor, situated learners gain knowledge through lived practice and by becoming a participant in a community of practice (Driscoll, 2005).

Games promote scientific practice through situated cognition. The game world becomes the community of practice, and the player is the apprentice. Squire (2006) stated that "in videogames, knowing is at its essence a kind of performance, as learners learn by doing, but within powerful constraints instantiated through software and social systems" (p. 26). In discussing an epistemic game called *Urban Science*, Shaffer et al. (2009) explained that "professional work and reflection on that work with peers and mentors in game form can be a powerful educational tool" (p. 4). Just like Crawford's (1984/2011) example of lion cubs at play, students playing professional roles within a game get to play, practice, and perform the skills necessary to do the job—free to learn through failure without the threat of a deadly mistake. Research has shown that when students take on professional roles in mobile AR games, they can engage in scientific inquiry (Dunleavy et al., 2009; Rosenbaum et al., 2007; Squire & Klopfer, 2007) and argumentation (Mathews et al., 2008; Squire & Jan, 2007) as they play. Researchers have even determined that students embodied their roles to such an extent that "students saw their abilities within the game as tied to their roles in authentic way, like a real occupation" (Rosenbaum et al., 2007, p. 40).

Game-Based Learning in K-12 Schools

Before a discussion of games can ensue, there is a need to define the term: *game*. According to Marc Prensky (2001), author of *Digital Game-Based Learning*, "the most basic definition of a game is that it is *organized play*, that is to say rule-based" (loc 2755). This definition bears striking resemblance to how earlier theorists denoted the connection between play and learning. In brief, educational games in school work best when they are actually games, rather than interactive worksheets disguised with cute animations (Klopfer, Osterweil, & Salen, 2009). Additionally, the idea of games and school are not mutually exclusive; in fact, a group of prominent researchers in games and learning, Klopfer et al. (2009) stated that "schools can and should play a critical role in fostering learning in association with game play" (p. 3). An entire curriculum can even by framed around game-based learning as exemplified by *Quest to Learn*, a New York City school (Salen, Torres, Wolozin, Rufo-Tepper, & Shapiro, 2011).

As discussed in the previous section, well-designed games encapsulate extremely effective learning theories. Thus, this section will briefly touch on some exemplars over the past few decades of game-based learning working effectively in classrooms. Game-based learning also has its skeptics (Clark, 2007; Healy, 1998); therefore, in order to present a balanced understanding of the knowledge landscape, this section will conclude with an in-depth discussion of the criticism of game-based learning.

Exemplars of Game-Based Learning in Schools

In 1981, computers games were essentially brand new and games immediately entered the classroom. A brief selection of exemplars from classroom gaming throughout the years will be discussed, starting from the early days of *Oregon Trail,* continuing with commercial-off-the-shelf (COTS) games that were designed for mass market yet repurposed for the classroom, and arriving at more recent game-based curricula designed with specific curricular objectives.

The early days. Adults today have fond recollections of playing *Oregon Trail* in school (Anderson, 2012). *Oregon Trail* was originally released in the earlier part of the 1980s, yet it is still around today and can be played within Facebook (Klopfer et al., 2009). It was one of the very first educational games, and certainly one of the most wellknown (Watson, 2010). While it is known that *Oregon Trail* has won many awards (Howell, 2010) and been incorporated into many classrooms (Watson, 2010), no real literature exists to understand the learning impact of the game. Researchers believe it was incorporated into classrooms because teachers could justify using it since it mapped easily to their history curriculum objectives (Klopfer, 2008). In other words, it opened the door to games in the classroom. In 1995, the NRC's commissioned report acknowledged the role that video games play in the culture of childhood and advocated for educators to examine the educational potential of video games. In the report titled—*Reinventing Schools: The Technology is Now!*—games in the classroom received strong support; the report noted that games had the motivational power to engage the Nintendo Generation (NRC, 1995).

Commercially available simulation games. In the 1990s, new COTS 'thinking games' came on the market and were not as easily incorporated into classrooms; they were inter-disciplinary and did not map easily to curriculum objectives (Klopfer, 2008). Although not as widely utilized as *Oregon Trail*, these new simulation games found success in smaller doses. Many scientists themselves use simulation techniques (Gee, 2003), so it is no surprise that simulation games have claimed some success in classrooms.

Simulations games such as *SimCity* and *Civilization* have produced important findings in classroom-based settings. In *SimCity*, players design and create their own cities; working with a budget, students learn their success and failure as they try out different urban plans. Kirriemuir and McFarlane (2004) reviewed many reports and, although the information is patchy, they consistently found favorable benefits of students using *SimCity* in classrooms—anything from math and engineering to urban planning and environmental awareness. Games such as these were also shown to enable group discussions and experimentation among the players. Successful experiences with such "simulation games could be, for some children, a precursor domain for those sciences that heavily trade in computer-based simulations as a method of inquiry" (Gee, 2003, p. 48)

The learning potential of *Civilization* has also been investigated in classrooms. In the game, players are challenged to guide a civilization from the Bronze Age through the Space Age, a span of 6,000 years! By playing the game, students are exposed to a lot of historical content, but more importantly, they explore socio-scientific and political issues (Clark, Nelson, Sengupta, & D'Angelo, 2009). When tested in the classroom, *Civilization III* provided evidence of a strong learning potential; it enabled student to think in terms of an approach to history, rather than history as a pile of facts (Squire, 2004). While some benefits are documented, researchers have also found that teachers struggle to keep kids on track and gaining the appropriate educational experience from playing such large-scale simulation games (Kirriemuir & McFarlane, 2004).

Educational games as designed curricula. Academic researchers are now crafting such games specifically for classroom use. More specifically, these new educational games are more like virtual worlds that leverage elements of multiplayer online role playing games (Barab, Sadler, Heiselt, Hickey, & Zuiker, 2010). McGonigal (2011) stated that "when these games work—when they marry good game design with strong educational content—they provide a welcome relief to students who otherwise feel underengaged in their daily school lives" (p. 128). Despite the criticism against game-based learning, research shows that these new game-based virtual worlds show promise, particularly in science: *Quest Atlantis* and *River City* are two well-researched exemplars of this new breed of classroom gaming.

After reviewing over 300 articles, researchers concluded that the academic value of the average science video game had little support; however, "*Quest Atlantis* provides the strongest evidence that science-based gaming can be useful for achieving deep learning" (Young et al., 2012, p. 72). Researchers concluded that the curriculum's effectiveness derived from the notion that "it is more consistent with how the content is experienced in the real world" (Barab et al., 2010, p. 405). Researchers even found evidence that student learning within the game-based curriculum transferred to standardized test items (Barab et al., 2007). There are two important ideas to take away from this project, the role-playing design element and the teacher support. First, Barab et al. (2007) concluded that "role-playing is a powerful design move for engaging students in meaningful interactions with designed environments for learning" (p. 779). Second, the game-based curriculum is a very well supported project; teachers who are going to use the curriculum participate in online professional development beforehand, and there is an online forum for ongoing support (NRC, 2011a). Teachers are even provided with tool kits and assessments (Watson, 2010).

Unlike *Quest Atlantis*, which is an online world that is available 24/7, *River City* is a fixed-duration online game; however, they both embed scientific processes within the game and enable students to use scientific tools to conduct scientific inquiry (Clark & Martinez-Garza, 2012). In *River City*, students take on scientific roles and work in groups (NRC, 2011a). Research showed that when students were guided to socially construct their knowledge in the game they had a stronger understanding of scientific inquiry than other students (Ketelhut, Nelson, Clarke, & Dede, 2010). Also, immersing students in professional roles seemed to impact their self-efficacy; researchers concluded that experiences such as *River City* could be used to increase player self-efficacy in science (Ketelhut, 2007). Ultimately, the findings demonstrated that students were engaged by the game-based curriculum (Ketelhut & Nelson, 2010) and students who exhibit certain neomillenial learning styles—such as creating and sharing artifacts online—were well suited for learning in the *River City* virtual world (Dieterle, 2009).

Using games in the classroom has been around since the early 1980s and there is a growing body of knowledge about the positive impacts of game-based learning: why are there not more games being used in today's classrooms? First, schools have had to adhere to the standards set by the No Child Left Behind Act of 2001. As Klopfer (2008) stated, "the final nail in the coffin of learning games was the No Child Left Behind Act of 2001" (p. 26). Teachers and parents express concern over using "games in lesson time since such skill development did not match the criteria assessed in high stakes national tests" (Kirriemuir & McFarlane, 2004, p. 19). The other big obstacle to using games for learning is the ever-present criticism against game-based learning in schools.

Criticism of Game-Based Learning in Schools

According to Nicholas Negroponte (1995), "most adults fail to see how children learn with electronic games" (p. 204); therefore the research needs to be convincing and concrete. Unfortunately, educational video games are a nascent field of research (Young et al., 2012) and the evidence about the effectiveness of games is inconclusive (NRC, 2011a). In order to draw some conclusions, researchers have reviewed the literature and conducted meta-analyses on several occasions; regrettably, the evidence is still inconclusive.

Positive findings. To understand an overview of what the field is saying, researchers sometimes use a meta-analysis to synthesize and re-analyze research that has already been done. In reviewing several meta-analytic studies, some researcher found positive findings. First, Vogel et al. (2006) conducted a quantitative meta-analysis with 32 game-based learning studies. The researchers found that regardless of the person or the situation, the cognitive outcomes were more dominant for games and simulations over traditional teaching. Second, Wouters, van Nimwegen, van Oostendorp, & van der Spek (2013) conducted a similar meta-analysis with 39 studies that focused on Serious Educational Games (SEGs). They reported that instruction with the games yielded higher learning gains that traditional teaching. This study also confirmed that students who played in groups learned more. Finally, Sitzmann (2011) used meta-analytic techniques to investigate 55 reports on computer-based simulation games and determined that game players scored higher on delayed measures of factual knowledge than those who acquired knowledge through more traditional means.

Negative findings. Meta-analytic techniques have also revealed negative findings and contradictory findings. In the early 1990s, Randel and Morris (1992) conducted a literature review to investigate the effectiveness of games over conventional instruction. Out of 67 studies, 38 studies revealed no difference between student performance and conventional instruction. Of the remaining studies with acceptable controls, 22 favored games while 3 favored conventional instruction. Some might argue that the educational games available now are much improved and would reveal different results. Unfortunately, Wouters et al. (2013) published a recent meta-analysis of serious games. While they did find evidence of learning gains from games, their findings also revealed that serious games are no more motivating than passive instruction at improving learning, a contradiction to the conclusion of Sitzmann (2011) that learning was enhanced by simulation games because of their intrinsic motivational power as contrasted with passive instruction.

Limited results. Inconclusive results is only the first criticism against the research results generated in the games and learning field; other criticisms include limited empirical evidence (NRC, 2011a) and limited research to support achievement gains. After reviewing over 300 articles, Young et al. (2012) concluded that within the current knowledge base there is an "absence of sufficient empirical research comparing knowledge gains from video games to different modes of instruction" (p. 62). Young et al. (2012) also concluded that there was limited evidence to support video games increasing student achievement in K-12 education. Studies that have measured achievement gains found no statistical significance (Annetta, Minogue, Holmes, & Cheng, 2009; Wrzesien & Alcañiz Raya, 2010).

Barriers to implementation. Critics of game-based learning will also highlight problems associated with integrating games into the formal school environment, including costs, teacher role, logistics, and curriculum alignment. First, in the past, the costs of licenses, software products, and video game equipment had to be considered and were thus a barrier to implementation (Kirriemuir & McFarlane, 2004). Second, the role of the teacher is complicated; students are not necessarily used to playing a game in a learning environment (Squire, 2005), so teachers need to consistently work at keeping kids on track towards the educational goal (Kirriemuir & McFarlane, 2004). Third, the constraints of the school day do not afford players several hours to immerse themselves in the game play and derive the maximum benefit (Young et al., 2012). Finally, there are always concerns from teachers and administrators that games do not map well to curriculum objectives (Kirriemuir & McFarlane, 2004).

Best Practices of Game-Based Learning in Science

Despite criticism against game-based learning and inconclusive research results, the big problems in education still exist and still need solutions. First, students are disengaged by school; there is a "recognized need to (re)engage disaffected learners, and game based learning is seen as a potentially effective response" (Arnab et al., 2012, p.163). Second, students need 21st century skills (Partnership for 21st Century Skills, 2013). As stated by David Williamson Shaffer (2006), "we can't skill and drill our way to innovation" (p.3); game-based learning offers an engaging educational paradigm that promotes the skills desperately lacking in today's workforce. Finally, readiness for the workplace is especially important for the job growth predicted in careers in science and technology (Lockard & Wolf, 2012). There is a growing body of research suggesting that Serious Educational Games can effectively promote scientific practices (Asbell-Clarke et al., 2012; NRC, 2011a; Steinkuehler & Duncan, 2008).

Since the turn of the century, researchers have noted a widening acceptance of the potential for game-based learning in the classroom, so the body of knowledge is starting to grow (Kirriemuir & McFarlane, 2004). Specifically, these games are Serious Educational Games, or SEGs for short. The term was originally used in 1970 by Clark C. Abt as the title for his book; now, it is generally used to describe the work in game-based learning since the turn of the century (Heneghan, 2008).

The body of evidence is growing and it's promising; Wouters et al. (2013) conducted a recent meta-analysis of the literature and reported that SEGs are superior to traditional instruction for both learning-in-the-moment and retention of knowledge. Vogel et al. (2006) also reviewed studies of games and simulations verses traditional teaching; they reported positive effect sizes for simulations and game-based learners in cognitive gains and attitude. However, the evidence is still slim for the impact of SEGs on student achievement (Young et al., 2012).

To make the best use of research to date, researchers need to cull together elements of best practices and apply them in an empirical study to measure student achievement. First, game-based learning shows benefits when the games target specific subject matter (Randel & Morris, 1992) and evidence is mounting that SEGs foster quality science learning experiences (Asbell-Clarke et al., 2012; NRC, 2011; Steinkuehler & Duncan, 2008; Young et al., 2012), so best practices will be drawn heavily from science learning games. This section will present a literature review on science SEGs and cull together best practices. Specifically, this section will highlight how SEGs for science learning can engage learners, encourage scientific practices, and promote collaboration.

SEGs Engage Learners and Promote Flow

Schools need to reengage presently disengaged students, but more importantly, engagement is a prerequisite for deep, meaningful learning. According to molecular biologist John Medina (2008), "better attention equals better learning" (p. 74). Capturing students' attention is important, but another component is essential as well. Csikszentmihalyi et al. (1993) contended that "we cannot expect our children to become truly educated until we…[know] how to spark the joy of learning" (p.195). Thus, the keys to meaningful learning in school are finding how to sustain attention and how to spark enjoyment: game-based learning can do both.

In February 2007, Sasha Barab and Chris Dede published the lead article in a special games and simulations issue of the *Journal of Science Education and Technology*. In discussing the issue as a whole, Barab and Dede (2007) emphasized that all the research published in that issue demonstrated that SEGs offer a powerful potential to engage learners in deep, meaningful, science learning. Through situated learning, well-designed SEGs immerse the player in a realistic experience that can promote engagement (Annetta, Murray, Laird, Bohr, & Park, 2006; Dede, 2009). Qualitatively researchers have confirmed improved student engagement while playing certain science learning games (Dunleavey et al., 2009; Gustafsson, Katzeff, & Bang, 2009). Researchers have also confirmed that participants' level of engagement had a statistically significant increase when playing a SEG (Annetta et al., 2009). According to Annetta and Cheng (2008), "educational games seem to be effective in enhancing motivation and increasing student interest in subject matter…[they] generate a much higher level of students' positive emotional engagement" (p. 2).

Researchers have commented on what aspects of games make them such engaging experiences. In 1981, Malone posited that the most important factors that make playing a computer game intrinsically motivating are challenge, curiosity, and control. More recently, researchers have found that player socialization is also connected to their engagement with the game (Barendregt & Bekker, 2011). Also, Hatfield (2011) argued that the best games provide not only challenging activities but also frequent and targeted feedback.

The mention of challenge, curiosity, feedback, and intrinsic motivation in relationship to game-based learning are pertinent to a discussion of flow, since they are all components of flow. A very small area of research is beginning to show evidence to support the notion that students playing SEGs are engaged, in part, because they are experiencing flow. Research from Bressler and Bodzin (2013) reported that the average player in their forensic science mystery game had a substantive flow-like experience. Other researchers have found confirmation of specific flow components such as time transformation (Wood, Griffiths, & Parke, 2007), positive affect (Wang, Khoo, Liu, & Divaharan, 2008), and motivation (Huang, 2011).

Researchers are beginning to collect evidence of players' flow experience in connection to specific game elements; such evidence can be used as best practices in future games. Wood et al. (2007) found that the games' complexity, use of multi-levels, missions, multiplayer interactions and narrative all were associated with player sense of time loss. Bressler and Bodzin (2013) determined that the participants' flow experiences seemed to stem from three experiential features: Participants felt a flash of intensity, a sense of discovery, and desire for higher performance. They attributed the experience of flow partially to the lack of technological frustration and ease of navigation associated with vision-based AR—using visual markers rather than locational triggers—as well as a game design element that shifted leadership evenly between group members.

SEGs Promote Scientific Practices

According to the K-12 science framework authored by the NRC (2012a), schools in the United States are not engaging students in authentic experiences of practicing science, and there is too much emphasis on discrete facts. To make the learning engaging, students need to feel the relevance and authenticity of the learning activity (Kirriemuir & McFarlane, 2004). As stated by Singh, Change, and Dika (2005), "instead of teaching science as an abstract undertaking, teachers need to bring science teaching to life by making real life connections and showing the usefulness of scientific knowledge" (p. 213). A report published by the NRC (2012a) asserted that students should emerge from their schooling with an understanding of science as a set of practices; they outlined relevant practices including asking questions, analyzing data, constructing explanation, and engaging in argument from evidence. Within these practices, there is a lot of overlap. Generally speaking, students should be evaluating ideas and constructing explanations from evidence; to do this effectively, they must practice and engage in argumentation (Reiser, Berland, & Kenyon, 2012). Game-based learning environments offer an excellent context for such practice. Nussbaum (2012) maintained that "argumentation is a type of language game in which both players make moves with words" (p. 114); therefore, scientific argumentation can seamlessly exist in a Serious Educational Game.

Players in multi-player games have a natural tendency to engage in scientific argumentation; in a study that examined online discussions in *World of Warcraft (WoW)*, researchers found that 86% of the discussions displayed "social knowledge construction" which included scientific discursive practices such as evidence-based argumentation (Steinkuehler &

Duncan, 2008). *WoW* was not designed explicitly to promote scientific practice, which speaks to the players' natural tendency to engage in argumentation. Games designed to support scientific practice yield similar results. Squire and Jan (2007) studied groups playing *Mad City Mystery* and observed that all groups engaged in argumentation. In a study with *Martian Boneyards*, researchers found evidence that the game nurtured sustained scientific inquiry during game play (Asbell-Clarke et al., 2012). Scientific practices that students engage in during game play also translate to performance based tasks. Dyads of student who experienced scientific practice in an immersive game world performed significantly better than control groups on a performance-based transfer task that required causal understanding (Barab et al., 2009).

Research shows that game-based learning in this context derives from the process of participating in a practice; the learning does not come from acquisition of facts (Kirriemuir & McFarlane, 2004). This research supports the Next Generation Science Standards, authored and published by Achieve, Inc. (2013) on behalf of the 26 states and partners that collaborated on the project, and based largely on the science framework published by the NRC (2012a). SEGs work well for promoting scientific practice because players engage in authentic practice by playing professional roles and working with scientific tools. David Williamson Shaffer (2006), author of *How Computer Games Help Children Learn*, begins his book with a quotation from Merlin Donald, a cognitive scientist: "What is a 'career' or a 'vocation' except a role-playing game extended over an adult lifetime." It's a very relevant quotation since Shaffer's book makes the point that the roles that the students play in these scientific role-playing games provide a means for thinking about the content in a relevant context (Shaffer, 2006). In a review of several, notable multi-user virtual environments, Nelson and Ketelhut (2007) confirmed that such "curricula can successfully support real-world inquiry practices based on authentic interactivity with simulated worlds and tools" (p. 265). Cheng and Tsai (2012) reviewed five studies from the field of location-based AR games and found that collaborative role-playing design encourages players to observe scientific phenomena, ask questions, investigate data, create hypotheses, and construct explanations—all important scientific practices.

SEGs Promote Collaborative Learning through Discourse

In order to solve the problems of tomorrow, students need critical thinking, problem solving, collaboration, and communication skills (Ramirez, 2013). The Framework for 21st Century Learning prepared by the Partnership for 21st Century Skills (2013) details several aspects of how students should be able to communicate clearly such as listening effectively and articulating thoughts. Research shows that effective group communication is a key predictor of group success (Barron, 2003; Pentland, 2012). Teaching these skills to students will ensure that they are well prepared for today's team-based work environment.

Students working in small collaborative groups have been a vision for progressive education since 1928 when Rugg and Shumaker published *The Child-Centered School*. Research has shown that when students work cooperatively in groups at the elementary level it can lead to significantly higher achievement in both reading and math over students in traditional settings (Stevens & Slavin, 1995). Working in groups is not only effective in lower grades but also in higher education. As mentioned by Thomas and Brown (2011), "A line of research begun by Harvard University professor Richard Light demonstrated that study groups dramatically increase the success of college students" (p. 67). Unfortunately, Demetriadis et al. (2012) summarized the relevant research on freely collaborating students and concluded that productive learning interactions do not occur when collaborative groups have no support or scaffolding. The reality is that—to learn how to collaborate effectively—students must be helped (Bransford et al., 2000).

Technology is one way to support collaborative learning. Collaboration with computers was first explored in the 1990s as an educational research topic. Researchers found that students trained in collaborative learning with educational technology had higher levels of achievement and self-esteem (Sivin-Kachala & Bialo, 1996). Honey and Henriquez (1993) found that working together on the computer helped students develop higher level thinking skills. Classroom research in the Apple Classrooms of Tomorrow project showed that students work effectively in groups when they used computers; it seemed to intensify the students' engagement and improve the quality of their produced artifacts (Fisher, Dwyer, & Yocam, 1996).

As the internet arose in the 1990s offering exciting opportunities for people to connect and learn together, a new branch of the learning sciences emerged (Stahl, Koschmann, & Suthers, 2006). Computer-supported collaborative learning, or CSCL, studies how people learn in groups with help from computer technology. According to Stahl et al. (2006), CSCL stresses that the "learning takes place largely through interactions among students. Students learn by expressing their questions, pursuing lines of inquiry together, teaching each other, and seeing how others are learning" (p. 410). Using the collaborative group as the unit of analysis, research in CSCL focuses on socially constructed knowledge gained through interaction. Bransford et al. (2000) summarized a body of research and concluded that collaborative group work increases the quality of the feedback that students receive thus increasing their knowledge and ability to revise their thinking. Songer (1996) reported that when students used the internet to dialogue with distributed peers, it seemed to increase motivation. CSCL researchers seek to support opportunities such as these and direct discourse to promote shared knowledge construction (Stahl et al., 2006).

Players in multi-player games have the capacity to work together in such a way that new knowledge is constructed. Using an online protein folding game called *Foldit*, gamers generated models that solved a problem that expert biochemists had failed to solve themselves (Khatib et al., 2011). Well-designed games foster effective collaborative learning and social construction of knowledge. As summarized by Kirriemuir and McFarlane (2004), “games are often a facilitator to social, communication and peer activities…learners gain resources from fellow members that equip them to solve problems within” (p. 15-16). SEGs can support collaborative discourse through positive interdependence and jigsaw pedagogy. According to Johnson, Johnson, and Holubec (1993), “positive interdependence is successfully structured when group members perceive that they are linked with each other in a way that one cannot succeed unless everyone succeeds” (p. 9). The notion of jigsaw pedagogy is that each student in a group becomes an expert on one aspect of the activity, and teaches it to the other group members (Aronson, 1978; Aronson and Patnoe, 2011).

The Promise of INPLACE Mobile Games

Experts foresee K-12 education incorporating game-based learning within two to three years (Johnson et al., 2012). As mentioned in the *2012 Horizon Report*, “the greatest potential of games for learning lies in their ability to foster collaboration and engage students deeply in the process of learning” (p. 6). The 2011 NRC report on games and simulations has also advocated for further investment in game-based learning, specifically as an avenue to improve science learning. They assert that games have potential to achieve many goals of science education including cultivating motivation to learn science as well as scientific argumentation and discourse. Overall, this chapter has built the case that well-designed games make for excellent learning environments, especially for skills that today's workplace demands. This section will define a new category of Serious Educational Games—called INPLACE mobile games—and explain how INPLACE mobile games address current criticism against game-based learning.

What Are INPLACE Mobile Games?

INPLACE mobile games combine the best practices of what the games and learning field knows about how to engage students, how to support collaboration, and how to promote authentic practice in a discipline.[3] INPLACE is an acronym created by this study's author that stands for: **I**nterdependent, **N**etworked, **P**articipatory **L**earning, **A**ugmented, **C**ollaborative **E**xperience. INPLACE mobile games began over a decade ago when the Massachusetts Institute of Technology (MIT) first conducted research on mobile AR games. Their first game, *Environment Detectives,* put students in groups to investigate a toxic spill using handhelds enabled with Global Positioning System (GPS); the project's purpose was to understand the learning affordances of AR games and demonstrate the project's potential for success in formal learning (Klopfer, Squire, & Jenkins, 2002). To move the research agenda forward, several universities collaborated on a large-scale implementation of an AR game called *Alien Contact* that was funded by the U.S. Department of Education. The research results showed high student engagement, but also logistical limitations (Dunleavy et al., 2009). Now with AR development platforms such as ARIS, TaleBlazer, and 7scenes that enable non-programmers to create their own INPLACE mobile games, the ability to make games no longer resides within the commercial sector but with educators and researchers; however, there are relatively few research

[3] While INPLACE games can be used for disciplines other than science, this study will be focused on science practices thus the definition will be built using research on mobile science games.

teams with a research agenda to understand how these games can enhance K-12 education (Dunleavy & Dede, 2013).

They are mobile. Mobility is a technological paradigm shift. As Laurillard (2009) stated, "the simple fact of its mobility has generated a new research field" (p. ix). Project Tomorrow (2010) surveyed thousands of students and concluded that "in the student vision, mobile devices have the potential to directly impact learning and personal productivity" (p. 9). Basically, when students learn effectively with mobile devices, their learning becomes contextualized, connected, and dispersed through time; all of these themes afford a more personalized learning experience for the student (Crompton, 2013).

They augment reality. Experts foresee K-12 education incorporating AR within four to five years (Johnson et al., 2012). Cheng and Tsai (2012) reviewed studies on AR in science education and concluded that most studies showed that learners had positive attitudes towards AR. Beyond just being 'the new thing' or 'something cool,' AR provides an important component of the immersive feel of the game. According to Squire and Jan (2007), these games augment reality because they are "games played in the real world with the support of digital devices (PDAs, cellphones) that create a fictional layer on top of the real world context" (p. 6). INPLACE mobile games use location-aware AR activated by GPS or visual markers such as quick-response codes (QR codes). This means that the games can provide location-specific information and players can experience content in context. Pedagogically, given the situated nature of the learning, the students are then better prepared to solve similar problems in the future because the learning transfer is more direct (Dunleavy & Dede, 2013).

They are networked. INPLACE games are *networked* in the technical sense of the word, as well as the social sense of the word. First, the games are supported by a network of computers.

While software might reside on the mobile device, it is the fact that the mobile device is connected to the internet that allows the game to distribute information to different players. Second, the games afford a rich network of interactions (Gee, 2003). Players are interacting with the storyline, with characters and tools, and players are interacting with each other in the real world. This fosters feeling of connection to the content and the teammates. At the Boston Museum of Science, Klopfer et al. (2005) found that *Mystery at the Museum* fostered feelings of connection between the visitors and the museum. Feedback from players confirmed that they were engaged, not only with individual exhibits, but also with the larger context of the museum and with their fellow players.

They encourage participatory learning. Fundamentally, INPLACE games shift learning from an acquisition model to a participatory model. Lave and Wenger (1991) described how students begin learning within a community of practice: "it crucially involves *participation* as a way of learning—of both absorbing and being absorbed in—the culture of practice" (p. 95). In the INPLACE games that have been developed and researched to date, the method of participatory learning that works is to have students assume roles within the game. For example, in *Outbreak @ The Institute*, players become doctors, medical technicians, and public health experts in order to determine to cause of a disease outbreak (Rosenbaum et al., 2007). The researchers determined that "students saw their abilities within the game as tied to their roles in an authentic way, like a real occupation" (Rosenbaum et al., 2007, p. 40). In *Alien Contact*, students became chemists, cryptologists, computer hackers, and FBI agents, in order to determine why aliens landed in the school yard (Dunleavy et al., 2009). Researchers reported that the students found role-playing very motivating. In fact, role-playing is so powerful that when Facer et al. (2004) had students break out of their roles to reflect on game strategy, they deemed it a failure of the game because it did not capitalize on students' self-motivation or conceptual understanding from the viewpoint of their roles. Dunleavy et al. (2009) even argued that their data showed students were experiencing *projective identity* as discussed by Gee (2003). When players identify so intimately with the roles, the games can leverage projective identity in order to create compelling, participatory learning experiences.

They are collaborative. Trespalacios, Chamberlin, and Gallagher (2011) found that middle school students have a strong preference for multiplayer games. However, INPLACE games are collaborative because it is not only the preference of the players but also a highly desirable way to learn that teaches them necessary skills. As discussed earlier, freely collaborating students often fail to engage in meaningful interactions; however, INPLACE games provide the scaffolding and support necessary so that students can work together effectively.

INPLACE games are designed to be collaborative, but the theory draws heavily from cooperative learning. In cooperative learning, students spilt the task into individually parts, work individually, and then assemble the final result, whereas, collaborative learning requires that students work together (Dillenbourg, 1999). Because of the design of INPLACE games, players exhibit both cooperative and collaborative learning. In order to scaffold this type of collaborative problem-solving, each player role's role is differentiated. According to Klopfer (2008), students playing collaborative mobile games "help each other, observe each other, and act together to create communities as they learn to solve problems" (p. 223). Scientific practice in inherently social; groups of people need to put their expertise together in order to solve problems. INPLACE games provide authentic, collaborative practice.

They are interdependent. Students not only work together in INPLACE games, their group success depends on the success of all the individuals. Researchers have found that interdependent roles are an effective way to scaffold collaborative problem solving within mobile learning games (Dunleavy et al., 2009; Squire & Jan, 2007). This can be done with jigsaw pedagogy in which students receive different information and have to teach other. For example, in *Time Lab 2100*, students played in pairs to travel back in time to makes change in an attempt to decrease the impact of climate change; each player only received half of the dialogue from characters who offered expertise (Klopfer & Sheldon, 2010). Another way to have groups work together is to differentiate the roles. In *Sick at South Beach*, students were water chemists, public health doctors, or wildlife ecologists; each student developed a different expertise over the course of the curriculum and argued quite intensely for their point of view when it came time to develop a consensus (Squire, 2010). Overall, the use of interdependent roles in INPLACE games scaffolds the social interaction and discourse necessary to build scientific knowledge offering an unmatched way for students to socially construct knowledge.

They are experiential. INPLACE mobile games have a kinesthetic component which offers an immersive feel—players are up and moving around. Researchers have found that this physicality contributed to players' enjoyment (Facer et al., 2004, Dunleavy et al., 2009). However, the immersive feel comes from more than just the physical experience. In reference to the capabilities of today's new media, Chris Dede (2009) explained that "immersion is the subjective impression that one is participating in a comprehensive, realistic experience" (p. 66). Several aspects of INPLACE games seem to support this description of immersion. First, narrative plays an important role; Dunleavy and Dede (2013) summarized the literature and declared that the narrative is a critical design decision. Second, given the portability of mobile devices, the real environment can also provide an authentic context. For example, students playing *Up River* became immersed in their local watershed; throughout gameplay, players developed a detailed understanding of the environment and the conservation issues surrounding it (Wagler & Mathews, 2012). Third, the combination of the authentic context and the INPLACE game can provide an extremely robust sense of immersion. In their *EcoMOBILE* project, Kamarainen, Metcalf, Grotzer, & Dede (2012) are using mobile AR interfaces to contextualize information and supply just-in-time instruction as students explore a real pond. Students use real probeware which enhances the realism because the technology is very similar to what real scientists use to take ecosystem measurements (Dede, Grotzer, Metcalf, & Kamarainen, 2011). For *EcoMOBILE*, being in the real context provides 3-dimensional experiential learning, but more notably, the mobile technology augments the experiential learning.

Moving Beyond Edutainment and Arriving INPLACE

INPLACE mobile games, a new category of Serious Educational Games, have promise and potential, even in light of the criticism against game-based learning. INPLACE mobile games offer something different—they capitalize on the affordances of the device, the best practices of game-based learning, and they try to help students learn content in authentic, participatory, collaborative ways. INPLACE games hold promise for K-12 education because:

1. *INPLACE games are a hybrid of entertainment and edutainment.*

Klopfer et al. (2009) explained that there are two camps for educational gaming: in one camp, the game is pinnacle and the education is lost, while in the other camp, the learning goals are the focus and the game design is lost. INPLACE games, with their focus on participatory learning and collaborative design, bring together the best of both camps—losing neither the game design nor the learning goals. The content is the game.

2. *INPLACE games are kinesthetic*.

When researchers have considered the problems with educational games, they have expressed concerned for the lack of kinesthetic activities to support the virtual learning, expressing concern that there is a loss of transfer from virtual reality to physical reality (Young et al., 2012). As expressed in the definition of INPLACE games, they offer an immersive experience within a real, authentic context. Kinesthetic activity is built right into the game design.

3. *INPLACE games require minimal costs*.

Critics have complained that school cannot afford the software or the equipment necessary to play games at school. The software necessary to play INPLACE games is mostly free. The platforms that are used to design INPLACE games are mostly freeware as well. Platforms such as ARIS (2013) are free to make games *and* free to play games. Critics may complain that the graphical nature of these platforms is not as realistic as what students are accustomed to outside of school. While that is certainly the case, Vogel et al. (2006) found no significant learning differences between games of higher and lower levels of realism. As far as equipment costs, INPLACE games are mostly designed for devices that students are already carrying with them, their mobile phones. The culture of school just needs to shift the paradigm from phone as a 'threat to learning' to phone as a 'support for learning.' In fact, letting students use their own devices can contribute to an even deeper personal learning experience.

4. *INPLACE games are designed for curriculum support*.

While teachers might complain that COTS games required a lot of time and support to align the game with curriculum goals, INPLACE mobile games are designed around content. Some games are even designed with interdisciplinary learning goals such as *Alien Contact* (Dunleavy et al., 2009).

Using INPLACE Games to Engage Middle Level Students in Scientific Practices

Today's students have the opportunity to learn in different ways, largely because of the technologies that surround them outside of school (Ito et al., 2010). One medium that dominates a young student's out-of-school time is gaming and it starts from an early age. In fact, "gaming represents the central form of early computer experience for kids" (Ito et al., 2010, p. 196). By middle school, students are high-frequency users of video games and web-based games (Spires, Lee, Turner, & Johnson, 2008). More importantly, students value the use of games for learning (Goyal, 2012; Project Tomorrow, 2010). Nikal Goyal (2012), who published his student perspective on the state of education, stated "I have learned so much more about math by playing games…than by hearing lectures and doing textbook problems" (loc. 5400); he goes on to clearly state that he wants to see games used in the classroom. He is not the only student voicing this opinion. Students in the *Speak Up* study—approximately 300,000 students—also advocated for games in the classroom because games can get them more engaged with the subject matter, help them to work in teams, and provide an easier way to understand difficult concepts (Project Tomorrow, 2010).

Middle Level Students

Middle school is a critical time to foster intrinsic interest in science and build students' self-confidence in their potential for success in science careers. Researchers have found that students' pre-high school science experiences play an important role in whether students pursue science-related careers (Tai, Liu, Maltese, & Fan, 2006). When students do not select careers in science, it creates a less scientifically-literate citizenship and ultimately impacts the nation's ability to thrive in the global marketplace. According to report commissioned by the NRC

(2007a), the innovative capacity of the United States relies on developing our children to participate in a scientific workforce.

Declining science interest is a particular problem for middle school girls; however, mobile AR games are showing promise for bolstering their interest in science. Research has shown that girls' interest in science decreases from age 12 to 16 (Lindahl, 2007). This decline is not surprising given that girls of this age report more anxiety about science along with less enjoyment of science (Desy, Peterson, & Brockman, 2011). In contrast, the pilot program of *School Scene Investigators* found that neither gender nor interest in science predicted variance in flow experience (Bressler & Bodzin, 2013). In fact, after gameplay, 93% of participants were interested in learning more about forensic science. Researchers concluded that the way participants experienced science within the game seemed to change players' perception of science which ultimately increased their interest.

Middle school is not only a critical time for offering an engaging scientific curriculum but also an intellectually appropriate time. The experimental intervention for this study will demand that students think logically and wrestle with ideas in hypothetical terms—notable behaviors of the formal operations period as defined by Piaget (Mooney, 2000). Since children do not experience formal operations until approximately 12 years old, this type of curriculum experiment would not work well prior to middle school. Specifically, Squire and Jan (2007) found that students in primary grades who played their mobile AR game *Mad City Mystery* were "less strong at comparing evidence and integrating them into an overarching narrative" (p. 22). In contrast, middle school students were able to piece together a coherent narrative based on collected evidence.

Promoting Engagement

In a study with 4000 middle school students, Spires et al. (2008) found that students want school to engage them and stimulate them. While researchers were not surprised by this finding, they were surprised that the students had clear perspectives about how this could be achieved. Basically, students wanted to use technology as a tool to learn in a project-based learning environment. Spires et al. (2008) concluded that "students see a clear link between the use of *technologies* in school and their *academic engagement*" (p. 511).

School Scene Investigators: The Case of the Stolen Score Sheets, a game first studied by Bressler and Bodzin (2013), was designed to engage middle school students by using iPhones as tools to solve a forensic science mystery. Players' engagement was assessed according to flow theory (Csikszentmihalyi, 1990). Findings from Bressler and Bodzin (2013) revealed that the game promoted a substantive flow-like experience for the average player. Specifically, participants' flow experience seemed to stem from three experiential features: Participants felt a flash of intensity, a sense of discovery, and desire for higher performance. Certain design choices seemed to promote these factors that are important to flow experience. The chase sequence, originally included to make the game feel realistic and fun, may have offered that flash of intensity. The mystery-based story arc may have afforded a sense of discovery. To solve the mystery, students were "looking for clues" and "figuring it out"—aspects of the game regularly mentioned when describing their favorite part. Certain parts of the game, such as decoding the locker combination, seemed to promote the right challenge-skill balance, one of the most highly encountered flow elements according to Bressler and Bodzin (2013). When students experience the right challenge-skill balance, they desire higher performance.

Additional design features aligned to other aspects of flow including clear goals, sense of control, and unambiguous feedback. Clear goals were presented to players as quests; when a quest popped up onscreen, it told a player where to go or what to do. Although mostly a linear game, there was some non-linearity built in to provide players some control over the order of game elements. To ensure that players experienced unambiguous feedback, the game ended in a "win state" so players knew they had successfully completed the game.

Researchers have demonstrated a connection between flow and engagement in games (Bressler and Bodzin, 2013), as well as a significant correlation between learners motivation and cognitive processing (Huang, 2011); investigating the connection between flow state and academic achievement would move the research agenda forward. In fact, Admirral et al. (2011) reported that "promising new directions for future research might focus on flow with game play as an important explanatory variable of student learning" (p. 1193).

Promoting Scientific Practices in Authentic Contexts

When Spires et al. (2008) conducted focus groups with middle school students, they found that students wanted and expected school to prepare them for future jobs. Specifically, the students wanted to use "technology in authentic 'real-world' contexts" (p. 509). According to Bressler and Bodzin (2013), one authentic context that students are already familiar with is their own school. School buildings and school property can offer an important and deeply personal context for the learner in which to create authentic science learning experiences.

In order to make *School Scene Investigators*—also known as *SSI*—feel like an authentic forensic science mystery within a school, the game began with the school's principal accusing the players of stealing the test answers to the standardized test. During gameplay, students collected forensic evidence and had to prove their innocence by determining who really stole the answer sheets. As they moved around the school, they encountered QR codes that they scanned to access game-related information; this interaction is shown in Figure 1. In some instances, a code revealed a conversation with a virtual character. For example, players encountered a student named Jane when they scanned the code located outside the library. At other times, scanning a code meant discovering evidence to keep in inventory. For instance, students retrieved of piece of crime scene evidence when they scanned the code at the principal's office.

Figure 1: Player scans a QR code during *SSI.*

Throughout the game, players were exposed to some basic elements of forensic science, like analyzing fingerprints, hair samples, and other trace evidence.

Forensic science was not part of the science curriculum, so *SSI* was an enrichment activity. The pilot study provided evidence that the game design promoted flow. Using the authentic context of the school environment, the game design can also promote scientific practices through participatory learning; however, "as educators, a challenge is to develop forms of game play in which players' productive practices can align with curricular goals" (Squire, 2010, p. 2569). In order to make a broader contribution to the literature, the next iteration of *SSI* needs to align to curricular goals.

In addition, INPLACE science games can be used to promote deeper science understandings by moving students beyond the lowest levels of Bloom's taxonomy (Bloom, 1984). Rosenbaum et al. (2007) found that the students did understand complex causality after participating in their authentic science activity using location-based AR technology. More investigations examining how game mechanics and social interactions contribute to learning and achievement gains is required (Sharritt, 2008; Young et al., 2012). In order to comprehensively investigate the learning process that unfolds during game play, a mixed methods approach is recommended (Cheng & Tsai, 2012). In the original *SSI*, students were collaboratively discussing the scientific evidence; by studying the collaborative discourse within the game and capturing a written artifact in both a control and an experimental group, research results can reveal how dialogue within the game contributes to learning.

Promoting Collaboration and Communication

Trespalacios et al. (2011) studied middle school students' video game preferences and found that students prefer multi-player games over single-player games. Specifically, the players preferred working in groups because it fostered companionship, collaborative, competition, and challenge. *SSI* was designed to promote collaborative problem solving and research results revealed that the game built better relationships among the players (Bressler & Bodzin, 2013).

Certain design choices such as jigsaw pedagogy and interdependent roles seemed to promote the positive impact of the collaboration. Briefly, the game design established small groups where no player had all the information necessary to solve the mystery. In other words, the group *had* to work together. Authentic reasons pertaining to the storyline established reasons for providing different pieces of information to different players. Additionally, not all players had quests at all times nor did they always have the same quest. This design decision encouraged a diversity of players to take a leadership role.

The use of interdependent roles provided a platform to establish reasons for separate players to have different information. In *SSI,* players can be the social networker, techie, science whiz, or photographer. Each student plays a role and collects a unique piece of evidence from the crime scene as well as each suspect. Essentially, each player becomes an ‘expert’ in a specific strand of evidence, such as blood, hair, fingerprints, or shoeprints.

The design of *SSI* also included a non-linear path through the game. Non-linearity afforded players a chance to have control over some decisions within the game—an important component of flow (Csikszentmihalyi, 1990). More importantly, O’Shea, Dede, and Cherian (2011) confirmed that incorporating a non-linear design improved student collaboration by diminishing the competition among groups.

By incorporating such game designs, INPLACE games rely on the social interactions among players as a key to the overall success of the games. Research shows that players are highly engaged because of the positive interdependence of their roles (Dunleavy et al., 2009). Students are also motivated by the authenticity of their roles as well as the communication and collaboration that the game requires (Rosenbaum et al., 2007). Educational research must begin to determine the role of social learning (Young et al., 2012) and INPLACE mobile games that incorporate multiple players are a viable test bed. Additionally, future research needs to make a better attempt at valid comparison groups (Clark, 2007). By studying the collaborative discourse within the game in both a control and an quasi-experimental group, research results from the next iteration of *SSI* can reveal how social interactions within the game contributes to players’ success.

The Promise of an INPLACE Mobile Game for Middle Level Science Education

INPLACE mobile games are not widely adopted yet in K-12 education, but they are proving to be very flexible and can be implemented in different settings. Some activities have been implemented at informal learning institutions: MIT worked with Boston Museum of Science and the Columbus Zoo (Coulter, Klopfer, Sheldon, & Perry, 2012) while the University of Wisconsin- Madison worked with the Minnesota Historical Society (Dikkers, 2012). Other projects are trying to blur the boundaries between formal and informal learning. Some game designs have incorporated some classroom-based learning and some outdoor experiences on the school grounds (Dunleavy et al., 2009; O'Shea et al, 2011). Harvard University and Radford University are developing an activity called *EcoMOBILE* that students use while on a field trip to a local pond enabling students to connect classroom learning to the real world (Kamarainen et al., 2012).

The research in such flexible environments is promising; now the field needs to understand how to incorporate such games into real school environments. In the formal setting, there is emerging research that the benefits of location-based AR games can be converted into vision-based AR inside a school building (Bressler & Bodzin, 2013). Vision-based AR uses visual markers, rather than GPS coordinates, to trigger virtual content. Previous research on the original *SSI* was conducted inside a school building and confirmed the capacity for students to experience flow during the game (Bressler & Bodzin, 2013). The research also revealed how the game design fostered effective collaboration and an authentic context for learning scientific practice. Given middle school students desire to be engaged, to use technology as a tool for learning in problem-based scenarios, and play games in the classroom—specifically multiplayer games—*SSI* is a promising model. Thus, as an essential part of the middle level enacted curriculum, an INPLACE mobile game based on the original *SSI* model may provide opportunities for middle level students to experience flow, use scientific practices, and collaborate successfully.

CHAPTER 3: METHODOLOGY

The purpose of this study was to determine if an INPLACE mobile game promotes flow experience, scientific practices, and effective team collaboration. This study investigated students participating in not only the game developed explicitly for this study but also the control activity implemented by the classroom teachers. This study employed an embedded mixed methods design (Creswell & Plano Clark, 2011). The primary strand was a post-test only control group design with a quasi-experimental intervention. Quantitative data collected after the intervention included: a self-report survey that empirically assessed flow and a written report that measured scientific practices. To investigate players' collaborative team experience, a qualitative strand was embedded during the intervention. This chapter presents the study's research questions, setting and sample, treatment and control, research design, instrumentation and data collection, and data analyses.

Research Questions

This study compared a control activity with an INPLACE mobile game on two dependent variables including flow experience and a culminating artifact that measured scientific practices. First, the study investigated:

RQ1: Do students in the INPLACE mobile game compared to the control activity report a more substantive flow-like experience and create culminating artifacts with more evidence of scientific practices?

a. Are there any differences between treatment groups for flow experience or culminating artifact score when investigated separately?
b. Do reading achievement scores influence any of the above relationships?

RQ2: Does flow experience predict culminating artifact score?

Second, the study qualitatively assessed the experience of student teams during both the INPLACE mobile game and the control activity.

RQ3: How do in-team communication responses during the INPLACE mobile game compare to in-team responses during the control activity?

RQ4: How do scientific practices of teams in the INPLACE mobile game compare to those of teams in the control activity?

RQ5: How else are treatment groups different when qualitatively analyzed at the team level?

Setting and Sample

This study took place in a large school district in the state of Pennsylvania. The intervention started on September 23, 2013 and concluded on September 27, 2013. A convenience sample of 269 eighth grade science students was obtained from one of four middle schools in the district. The school is located in a diverse, urban area with mostly low-income households. Over one-third of the student population (35%) participates in the free and reduced lunch program.

Participants

To investigate the potential for conducting research, the researcher met with teachers and administrative staff of the school. To achieve a representative sample of students, eighth grade science teachers who taught a variety of track levels and at least four class periods were requested to participate in the study; they also needed at least one year of teaching experience. Two eighth grade science teachers agreed to participate in the research. Each teacher was responsible for five class periods of science. Shown below in Table 1 are the number of participants from each period and teacher. Including all 10 class periods in the study yielded a convenience sample of 269 participants.

Table 1

Original Participants by Period and Teacher

	Teacher A	Teacher B
Period 1	N = 19	N = 27
Period 2	N = 26	N = 27
Period 5	N = 27	N = 22
Period 7	N = 30	N = 30
Period 8	N = 29	N = 32
TOTAL	N = 131	N = 138

Sampling Design

As a mixed methods study, there were two different datasets collected. The quantitative dataset focused on the student as the unit of analysis and employed random assignment. The qualitative dataset focused on the student team as the unit of analysis and employed "purposeful random sampling" (Patton, 2002, p. 240). Sample assignment procedures for each dataset are discussed below.

Random assignment. For the quantitative dataset, the student was the unit of analysis. However, individual students were not randomly assigned to the experiment or control group for two reasons: the activities required collaborative teams and having both activities running in the same class was undesirable (e.g., students would notice some teams were playing a game and this might impact their behavior). Therefore, random assignment occurred in two steps, assignment of whole classes to treatment groups and then assignment of students to teams.

Assignment of whole classes. There were 10 classes in the study; all classes were either assigned as a control group or experimental group. First, to ensure an even distribution of achievement levels between the groups, classes were ranked according to the average student's mathematics score on the Pennsylvania System of School Assessment (PSSA). In other words, by using PSSA scores as a guideline each teacher ranked their classes on a scale of 1 to 5, where 1 was the lowest achievement level and 5 was the highest level. Second, for each rank, classes were randomly assigned to either the experiment or control group. Since the experiment required iPads and there was only one class set, a minor adjustment was made to ensure that the experiment was not assigned twice during the same period. Please refer to Table 2 for the final rank order and treatment assignment of each class period.

Table 2

Rank Order and Treatment Group for Whole Classes

	Teacher A	Teacher B
Period 1	Rank 1 Experiment	Rank 5 Control
Period 2	Rank 2 Control	Rank 4 Experiment
Period 5	Rank 5 Experiment	Rank 1 Control
Period 7	Rank 4 Control	Rank 2 Experiment
Period 8	Rank 3 Experiment	Rank 3 Control

Assignment of students to teams. The second step of random assignment assigned individuals to teams. In both treatments, students worked collaboratively in teams of three or four depending on class size. Using class lists and a table of random numbers, students were randomly assigned to teams.

Purposeful random selection. The process of team selection for the qualitative strand was purposeful random sampling (Patton, 2002). In order to identify important common and contrasting patterns, teams were purposefully selected in order to achieve a continuum of achievement levels (above average, average, and below average) and representation from both groups (experiment and control). In order to increase the credibility of the findings, one team was randomly selected to satisfy each category. As shown in Table 3, there were a total of six teams selected. Each one represented a case to be studied.

Table 3

Purposeful Random Selection of Student Teams

	Above Average	Average	Below Average
Control	Team C1	Team C2	Team C3
Experiment	Team E1	Team E2	Team E3

Treatment and Control

In order to explain the research design, a brief understanding of the treatment and control activities will be described in this section. The design of the quasi-experimental intervention drew heavily on the game design of *School Scene Investigators*, a game first studied by Bressler and Bodzin (2013). *School Scene Investigators: The Case of the Stolen Score Sheets* was a forensic science mystery game played on iOS devices that utilized QR codes in a school environment. Bressler and Bodzin (2013) found that the game built better relationships among players and promoted a substantive flow-like experience for the average player.

For this study, a new game was designed to build on the prior work of Bressler and Bodzin (2013). The new INPLACE mobile game, *School Scene Investigators: The Case of the Mystery Powder,* was the quasi-experimental intervention (see Appendix A for a brief game description). In order to design the game to promote collaboration and flow experience, design decisions from the original game were incorporated into the new game. For details about the game design of the original *School Scene Investigators*, please refer to the literature review in

Chapter 2. Since the original game was built with ARIS, the new game was also built with the ARIS platform (ARIS, 2013); this open-source, web-based programming environment offers a free iOS application and allows anyone to freely create and play games. The game relies heavily on QR codes; however, students were prompted to conduct hands-on experiments and insert information about their results into the game. Each student used a school-owned iPad to play the game; they played in teams. Student outcomes after the game were compared to student outcomes after participation in the control activity.

The control activity for this study was the mystery powder lab activity, a pre-existing curriculum unit adopted by the school district. Conducted early in the 8th grade school year over the course of three to five days, the mystery powder lab activity exposes students to the scientific method and promotes a basic understanding of the nature of matter (see Appendix B for the original documentation). In concert with the teachers and the principal, it was selected as the control activity for several reasons:

- Students engage in scientific practices described by the National Research Council (2012a).
- It is implemented as a collaborative scientific investigation with small groups of students.
- It has the element of mystery.
- It has already been taught for at least one school year.
- The content lends itself to game-based learning.

Research Design

This study employed an embedded mixed methods design (Creswell & Plano Clark, 2011). The primary strand was a post-test only control group design with a quasi-experimental intervention. A qualitative strand was embedded during the intervention to thoroughly understand participants' collaborative team experience in both the control and experiment. The quantitative strand investigated whether the two groups were different *after* the intervention while the qualitative strand investigated whether the two groups were different *during* the intervention.

The embedded mixed methods design was chosen because it balances the strengths and weaknesses of using randomized control trials (RCT) alone. Although RCTs are considered a powerful evaluation method, critics of RCTs have claimed that "the experimental design relies on the use of a linear understanding of causality" (Creswell & Plano Clark, 2011, p. 359). In other words, to say that any difference between the experimental and control groups is due to the treatment is a very reductionist approach; other factors, particularly social factors, could be causing changes to occur during the treatment. By embedding a qualitative strand during the experiment, attention can be paid to the social process of learning, not just the quantitative product.

When using RCTs in educational research, there is always a threat that the control group may receive treatment from the teachers that contaminates the quasi-experimental intervention. Diffusion of treatment could be a threat because a teacher may start using some of the pedagogical prompts from the game. To minimize this risk, the researcher had several meetings with the participating teachers prior to the research trial to remind them of the importance of teaching the control activity exactly the same as they had done in previous years. Two researchers were also onsite during research trials to facilitate data collection and offered similar reminders during the trials when necessary. Researchers were also cognizant of the fact that control groups may feel "resentful demoralization" (Shadish, Cook, & Campbell, 2002, p. 80) over not being included in the experiment. Neither researcher noticed any manifestation of anger or resentment. Some control students asked what the other classes were doing on the iPads but such students never exemplified resentment, just simple curiosity.

Instrumentation and Data Collection

As explained above, all classes in the sample were assigned to the control group or the experimental group; all students within each class took part in the appropriate activity.

Sample Size and Attrition

All participants (n = 269) produced a culminating artifact to measure scientific practices. Unfortunately, due to absenteeism and lack of class time on the last day, a subset of students (n = 35) did not complete the flow survey. Some participants were missing PSSA data for various reasons including having just moved to the district, this reduced the sample by another subset (n=12). Finally, an additional subset of students (n = 43) had to be dropped from the sample because a student walkathon during the middle of the intervention completely interfered with the experimental experience; the loss of class time meant moving all QR codes into the classroom and not enough time to finish the culminating artifact.

Table 4

Final Participants by Period and Teacher

	Teacher A	Teacher B
Period 1	N = 8	N = 24
Period 2	N = 24	N = 20
Period 5	N = 27	N = 15
Period 7	N = 27	N = 4
Period 8	N = 0	N = 30
TOTAL	N = 86	N = 93

As shown in Table 4, the final sample size was 179 students; all students had PSSA scores and participated in both post-treatment measures, including a post-survey to empirically assess flow and a culminating artifact to measure scientific practices. The sample population was comprised of all eighth grade science students and included 48.6% male (n=87) and 51.4% female (n=92) students. In addition to the full sample, a small sub-sample of 6 student teams, comprising 23 students, participated in the qualitative data collection. The alignment of data collection procedures with research design strands is shown in Figure 2.

	Quantitative Strand	**Qualitative Strand**
Individuals **(N=179)**	Flow experience survey; culminating artifact, PSSA scores	Not applicable
Teams **(N=6)**	Not applicable	Audio recordings of collaborative discourse, photographs, field notes

Figure 2: Matrix of research design and data collection.

Post-Survey

Immediately after completing the activity, participants completed the post-survey (see Appendix C). The post-survey was designed to measure individual demographics and flow assessment. The first part assessed each student's flow experience during the activity. The survey included 10 Likert-style items that measured elements of flow as outlined by Csikszentmihalyi (1996). Possible scores ranged from 10 to 50. The survey was derived from the Short Flow State scale (S FSS-2) and the Core Flow State scale (C FSS-2) developed by Jackson, Eklund, and Martin (2010). The survey was first used by Bressler and Bodzin (2013); Cronbach's alpha for that study was .80. After students registered their responses to the flow experience items, students indicated their gender, team composition, and whether they felt more curious about forensics.

Culminating Artifact

The culminating artifact for this study was an authentic assessment. This study did not use a traditional post-test assessment with multiple choice items because research showed that open-ended explanations captures student's knowledge integration ability better than multiple-choice items (Hee-Sun, Liu, & Linn, 2011). Furthermore, in order to measure the deeper learning as outlined by the NRC's Framework for K-12 Science Education, new assessment systems need to be developed (NRC, 2012a). Finally, according to Ketelhut et al. (2010), a standardized content test does not accurately measure the scientific practices gained during gameplay. Given this research, the learning outcome from this study was measured with an authentic assessment: a written report requiring open-ended explanations.

As part of both treatments, participants wrote a report as their culminating artifact. Students in the control group received an instruction sheet (see Appendix D) to facilitate completing the artifact which was called an *Investigation Sheet* (see Appendix E). To stay in line with the context of the game, and to replicate an authentic experience, students in the experimental group wrote an *Incident Report* (see Appendix F); no instruction sheet was provided because the game included all the prompts necessary to complete the incident report. These measures asked the same questions about scientific practice; the only differences were their titles and some minor wording adjustments to stay in line with game terminology.

Briefly, the culminating artifact contained a description of the incident or experiment that was being investigated, an investigation plan outlined as steps, known powder observational data, a known powders data summary, testable hypotheses for the mystery powder experiment, mystery powder observational data, an explanation of what each test revealed about the mystery powder, and a concluding paragraph that explained the compounds present in the mystery powder using supporting evidence.

Qualitative Data Collection

In order to investigate how teams in the quasi-experiment differed *during* the intervention from teams in the control, qualitative data was collected from six teams. During the entirety of the intervention, selected teams were audio recorded as well as documented with photographs and field notes. Onsite researchers took a series of photographs to document student interactions within the experiment and the control groups on all days. Field notes included observations of each period along with informal interviews with the teachers.

Audio recordings captured student discussions of selected teams for the entirety of the intervention. For the control teams, two audio recording devices were placed in the center of the table and recorded data from the beginning to the end of each class period. Two recorders meant that there was a backup recording in case something failed. For teams in the experiment, recordings were conducted at the individual level. Every participant on the student team wore a lapel microphone attached to a small digital audio recording device placed inside a pocket. To ensure high-fidelity of the qualitative data, all collaborative discourse was audio recorded and transcribed. Transcription was completed by DBS Associates in Princeton, NJ. The transcriptions document the conversation of each team on a daily basis. Unique speakers were not identified; however, turn-taking was clearly delineated.

Procedures

The principal of the school agreed to the proposed project. For a copy of the support letter, see Appendix G. Students in all classes taught by the participating teachers were expected to participate in the research study since the activities were approved by the superintendent and part of the adopted curriculum. Both student assent and informed consent forms were distributed to participating classes at the school in August 2013, and then returned to the researcher in order to use the student data for the study. The student assent form asked for the student's signature (see Appendix H). Informed consent was acquired from legal guardians or parents (see Appendix I for form).

Prior to data collection, the primary researcher built and tested the quasi-experimental intervention (also known as the *game*) in the school. Data collection began on September 23, 2013 and concluded on September 27, 2013. Some class periods finished the activity in four days; for others it took five days. Qualitative data was collected every day that the students participated in the activity. The flow assessment measure was given at the conclusion of the activity while the culminating artifact was integrated through the activity and written in parts on a daily basis until it was fully completed at the conclusion of the activity.

Post-Survey

Data from the post-survey was collected immediately after the activity concluded. For students in the control group, they were provided with laptops and directed to the survey URL hosted by SurveyMonkey (1999-2014), a leading provider of web-based survey solutions. For students in the experimental group, the end of the game automatically directed the students to the survey URL. These students filled out the survey from an iPad. In order to maintain anonymity of the individuals in the sample, each student used their student identification number—a unique numerical identifier—with the survey answers.

Culminating Artifact

Participants created their culminating artifact over the course of the activity. Students were given the report forms during the first day of the activity. To guarantee that the artifact was completed, teachers ensured that students were filling out each section as they progressed through the activity. In other words, students were not allowed to move on to the next phase of the activity until they wrote information in their report. Again, to maintain anonymity, students wrote their student identification number on the report form. Teachers were always available to answer questions. In some cases, the teacher provided scaffolding to assist students who were struggling but they did not directly provide the answers. Written reports represent each individual student's work. Some students were caught copying from a fellow student in the team; however, the behavior was minimal and existed equally in both the experiment and control groups. Original reports were maintained by the research team for grading, while participating teachers were given copies to use for their own purposes.

Qualitative Data Collection

Two researchers were onsite during the entirety of the intervention. One researcher was responsible for collecting qualitative data from the control classes, while the other researcher was responsible for collecting such data from the experiment classes. At the beginning of each class period, the researcher distributed the audio recording devices to the appropriate students. The researcher started the recording function on each device and tested the audio inputs. In the control group, devices were placed on the team's table. In the experiment, devices were given to team members to wear. During the period, the researchers observed and took pictures. In the experimental group, the researcher did assist with some iPad issues because she was known to the class as a technology assistant rather than a researcher. At the end of the period, the researchers collected the devices and stopped the recordings. Researchers discussed observations during free periods and at the end of the day. Field notes were written at these times.

Data Analysis

As an embedded mixed methods design, the researcher analyzed the primary data set (quantitative) first using statistical models to answer the primary research questions. Then, an analysis of the secondary data set (qualitative) was conducted using a descriptive case study approach to answer the secondary research questions (Miles & Huberman, 1994; Yin, 2014).

Primary Analysis

Before quantitative analysis began, the data needed preparation: the calculation of *total flow experience score (FLOW)* and *total report score (REPORT)*. Using the data from the post-survey instrument, total flow experience was calculated. Each level of the Likert-scale had a different numeric value with “I strongly disagree” equal to 1 and “I strongly agree” equal to 5. Since it was a 10 question survey, possible score range was 10 to 50. Cronbach’s alpha for this instrument was .77.

To calculate total report score, two raters scored all reports using a rubric (see Appendix J). The rubric enabled grading of written statements on a Likert-style scale where “insufficient response” was equal to 0 and “exemplary response” was equal to 4. Students were graded on six of the eight scientific practices as outlined by the NRC (2012a). The six practices were:

- Practice #1: Asking questions and defining problems
- Practice #3: Planning and carrying out investigations
- Practice #4: Analyzing and interpreting data
- Practice #6: Constructing explanations
- Practice #7: Engaging in argument from evidence
- Practice #8: Obtaining, evaluating, and communicating information

The instrument contained individual items pertaining to each practice. For example, for *obtaining, evaluating, and communicating information*, students were rated on how well they organized the information throughout the report and how well they explained ideas in writing. Students were rated on a total of 9 items; possible score range is 0 to 36. For examples of actual student responses, refer to Appendix K.

To begin the process of scoring the reports, three coders informally tested an initial draft of the grading rubric. Each coder independently graded four student reports representing a range of achievement levels. Disagreements were discussed and the grading rubric was revised. The process was then repeated. After the second set of four reports, the instrument was deemed reliable and formal coding commenced. A pilot test was conducted by two of the original coders. A class set of reports was randomly picked; it contained 30 reports. Coders independently graded the set and had an initial inter-rater reliability of Pearson's $r = .96$. Pearson's r was chosen since the scale was continuous, responses had a rank order, and there were only two raters (Leech, Barrett, & Morgan, 2008). Reasons for mismatches were primarily due to one coder's lack of scientific knowledge and a lack of familiarity with the writing capability of eighth graders. The reliability was deemed strong and both coders graded the remaining set of reports ($n = 239$). Final inter-rater reliability was Pearson's $r = .90$ for the complete set of 269 reports. The two coders met for several extended sessions to discuss all mismatches until unanimous consensus was eventually achieved. Reasons for mismatches were again due to one coder's lack of scientific knowledge while some discrepancies existed because one coder incorrectly applied the rubric.

To answer the first research question, a 1 by 2 factorial multivariate analysis of covariance (MANCOVA) was implemented and analyzed using SPSS. The MANCOVA analysis was chosen because there are two continuous dependent variables: total flow experience score (FLOW) and total report score (REPORT). While separate univariate analyses were possible, the MANCOVA analysis reduced the chance for Type I error and took into account the relationship among the dependent variables.

The categorical variable in the analysis was *Treatment Group* and it had two categories: experiment and control. Since intact classes were assigned to groups and the school district uses math PSSA scores to determine students' track level in school, math PSSA scores were used as a covariate to eliminate systematic bias and make group comparisons fairer. The scores were obtained from the school principal and represent test scores from the end of the seventh grade.

The MANCOVA analysis was used to evaluate the amount of variance in flow experience and report score as explained by treatment group. The primary hypothesis for the first question is that the experimental group will show higher levels of flow experience and higher scores on the culminating artifact; the MANCOVA tested for this as multivariate significance. To answer question 1a, univariate tests were conducted to determine which variables (FLOW or REPORT) were contributing to the multivariate significance.

To investigate question 1b, a 1 by 2 factorial multivariate analysis of covariance (MANCOVA) was implemented and analyzed using SPSS. The control activity required very little reading because students simply followed basic directions and instructional prompts (see Appendix D). Yet, as shown in Figure 3, the experimental activity demanded a high level of reading comprehension to participate in the activity. Players needed to read in order to understand the details of their inventory items, the instructional prompts, and the text-based conversations with characters. Without strong reading comprehension, the game is difficult to play. Additionally, the completion of the culminating artifact required a high level of writing ability. Overall, reading and writing ability may have influenced the relationships analyzed within the original MANCOVA. Thus, to answer this research question, the original MANCOVA analysis was re-run with reading PSSA scores as the covariate.

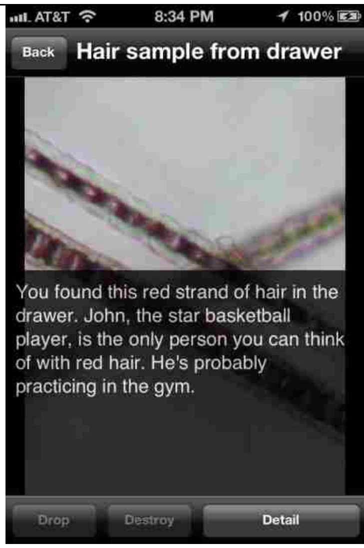
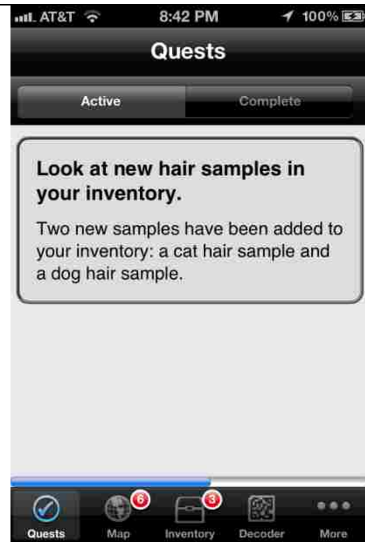
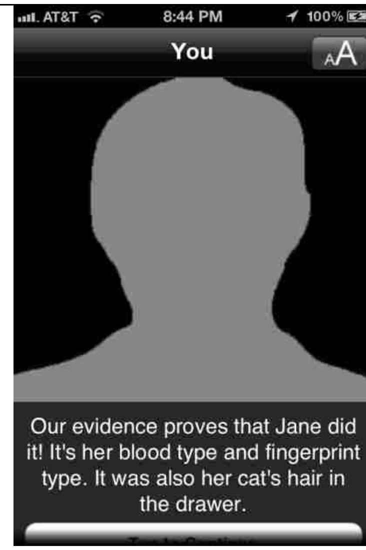
Details of inventory items	**Instructional prompts**	**Text-based conversations**
		

Figure 3: Samples of reading required during gameplay.

Finally, for the second research question, a multiple hierarchical regression model determined if flow experience (FLOW) predicted culminating artifact score (REPORT). Using SPSS, flow experience score was used to predict report score. In order to control for math PSSA scores and treatment group affiliation, they were entered into Block 1 of the regression model; flow was entered as Block 2.

Secondary Analysis

A secondary analysis was conducted incorporating a descriptive approach using multiple case studies with student teams as the unit of analysis (Miles & Huberman, 1994; Yin, 2014). According to Yin (2014), “the more that your questions seek to explain some present circumstance (e.g., ‘how’ or ‘why’ some social phenomenon works), the more that case study research will be relevant" (p. 4). Since the research questions stem from understanding the differences in the social process of learning within teams from different groups, case study research was chosen as the analysis method. Audio transcripts, photographic evidence, student reports, and field notes were compiled for within-case and cross-case analysis.

Coding strategy. Overall, there were six conversational transcripts—one for each student team—representing four to five class periods worth of team discussions. To analyze this data, the researcher chose two contrasting but complimentary strategies: (1) relying on the theoretical foundation and (2) working with the data from the "ground up" (Yin, 2014). To implement these strategies, there were two levels of coding. The first level was a priori based on the literature review, while the second level was emergent coding based on close reading of the transcripts.

For the a priori coding, two different sets of predetermined codes were used (see Appendix L for code lists). The first set of codes enabled the investigation of the third research question: How do in-team communication responses during the INPLACE mobile game compare to in-team responses during the control activity? Barron (2003) found that successful collaborative teams offer significantly more engaged responses than less successful teams. Building on the work of Barron (2003), this research analysis utilized her code structure of (a) accept, (b) discuss, and (c) reject. When students accept or discuss a teammate's proposal, then they are offering engaged responses. When students reject another teammate's proposal, then they are offering non-engaged responses.

The second set of predetermined codes facilitated the investigation of the fourth research question: How do scientific practices of teams in the INPLACE mobile game compare to those of teams in the control activity? Squire and Jan (2007) determined that mobile AR games can promote scientific argumentation; their codes included question, hypothesis, counter-hypothesis, and evidence. This research study expanded on their code structure and utilized codes that aligned directly to the scientific practices from the science framework published by the NRC (2012a).

To answer the fifth research question, a second level of emergent coding was required. Working with the data from the ground up enabled the investigation of the final research question: How else are treatment groups different when qualitatively analyzed at the team level? During a close reading of all transcripts, the researcher created memos with ideas for new themes or concepts. Then, the researcher developed several new code categories to further interpret team interactions within both treatment groups. In order to explain the patterns of engaged responses among treatment groups, codes were created for *commands* and *communal language*. While reading the transcripts, the researcher noticed a prevalence of commands ("Put more [water].") within control group while there was a prevalence of communal language in the experimental group ("Anybody have any quests? Okay. Then let's go.") While reading certain transcripts, the researcher noticed a large amount of confusion as well as off-topic conversations. Codes were then created for *confusion* and *off-topic discussions*. All transcripts were then coded again using the new codes. ATLAS.ti was used to code all the data.

Within-case analysis. Beginning with the resulting code reports, the author wrote a case story as a chronologically ordered narrative for each student team incorporating details from transcripts, the photographic evidence, the students' reports, and the observational field notes. Multiple sources of evidence were included to ensure triangulation and strengthen the validity of the findings. In the case stories, the researcher discussed teacher involvement and how the teams differentiated their roles and responsibilities within the group. The case stories were then condensed through a series of steps. First, each class period was summarized as a word table to document prevalence of concepts such as scientific practices, engaged responses, role differentiation, and teacher involvement. Second, the larger chart was collapsed into one final within-case display for each student team. According to Miles and Huberman (1994), "these displays can supply the basic material for explanations—plausible reasons for why things are happening as they are" (p. 90). The chosen display format was a conceptually clustered matrix (Miles & Huberman, 1994); the columns represented variables and concepts that aligned to the research questions, and rows represented cases. Finally, each case story was reduced to a short overview to provide a perspective of how prevalent each concept was within each student team.

To establish the accuracy and credibility of within-case findings, researchers maintained prolonged engagement with the cases under study and then participated in informal and formal peer debriefings. Researchers were present to observe the entirety of the intervention and record data. Formal peer debriefings included a written record that could be consulted later. The purpose of the debriefings was for the peer researcher to probe for biases, explore alternate meanings, and clarify interpretations of the author/researcher (Lincoln & Guba, 1985). Based on peer feedback, the researcher adjusted the case stories and data summaries.

Cross-case analysis. In order to enhance generalizability and to deepen the understanding of the team experience within each treatment group, a cross-case synthesis was conducted (Miles & Huberman, 1994; Yin, 2014). The researcher noted similarities and differences along the way; however, the formal cross-case synthesis did not take place until all six within-case analyses were completed. To permit systematic comparison, the case level displays were stacked into a meta-matrix display (Miles & Huberman, 1994). Specifically, a case-ordered descriptive matrix was chosen with the control cases placed in the top three rows and ordered from high achievement to low achievement and the experimental cases placed in the bottom three rows also ordered according to the same descending achievement levels.

To conduct the cross-cases synthesis, a pattern-matching procedure was utilized to examine the prevalence of concepts in all of the control cases, and then in all experimental cases, and finally in comparison to each other (Yin, 2014). In other words, findings were generated by making comparisons between cases within a treatment group, but making contrasts across treatment groups. A synopsis was written for each finding and supported with specific quotations and observations from the intervention; photographic evidence was included when possible.

To confirm all findings, the researcher underwent an external audit. According to Lincoln and Guba (1985), the external audit is “the most important step in providing evidence of dependability and confirmability” (p. 243). The auditor was involved after the analysis was essentially completed. In order to attest to the dependability of the analysis, the auditor examined the process of inquiry as well as the product. The data, the findings, and the conclusions were all part of the audit in order to establish confirmability of the findings.

CHAPTER 4: DATA ANALYSIS AND FINDINGS

This study investigated if an INPLACE mobile game promoted flow experience, scientific practices, and effective team collaboration with eighth-grade students. The game was compared to a traditional hands-on lab activity in which students worked in small teams. Recall from chapter three that the study utilized an embedded mixed methods approach with the primary quantitative analysis conducted with a sample of 179 students and a secondary qualitative analysis conducted with six student teams.

Data sources and instruments included (1) student flow scores from the post-survey (FLOW); (2) student report scores from the culminating artifact (REPORT); (3) math and reading PSSA scores; (4) audio transcripts; (5) photographic evidence; and (5) field notes. Statistical analysis of the quantitative data was conducted with *IBM SPSS Statistics for Windows Version 21.0* (IBM Corp., 2012) and qualitative coding and analysis were conducted with *ATLAS.ti* (2013). The subsections on subsequent pages present the findings from the primary quantitative analyses followed by the secondary qualitative analyses.

Primary Analysis

This section contains the MANCOVA analyses that used flow experience scores and report scores of individual students as the dependent variables and treatment group as the independent variable. It also includes a regression analysis investigating flow experience scores as a predictor of report scores. First, a MANCOVA analysis using math PSSA score as a covariate investigated multivariate significance between treatment groups. Second, a MANCOVA analysis using reading PSSA score as a covariate investigated multivariate significance between treatment groups. Finally, a multiple hierarchical regression analysis investigated flow as a predictor of report score while controlling for math PSSA scores and treatment group.

Descriptive Statistics

Before quantitative analysis began, the raw data was used to calculate the outcome variables: *total flow experience score* (FLOW) and *total report score* (REPORT). Descriptive statistics are reported in Table 5.

Table 5

Descriptive Statistics for Outcome Variables

Variable	M (*n*)	SD	Possible Range
FLOW			10.00-50.00
Experiment	40.51 (59)	6.62	
Control	38.62 (120)	4.53	
Total	39.24 (179)	5.37	
REPORT			0.00-36.00
Experiment	25.41 (59)	4.78	
Control	22.63 (120)	5.53	
Total	23.54 (179)	5.44	

Assumption Testing

Before using the MANCOVA analysis, certain assumptions were tested and met. The researcher checked for multivariate normality, homogeneity of regression, and the significance of regression. First, several steps were taken to check for multivariate normality. To begin the check for multivariate normality, the researcher established the univariate normality of both dependent variables: FLOW and REPORT. According to Lomax (2001), for univariate normality to exist, skewness and kurtosis must be between -2 and +2. Flow had a skewness statistic of -0.086 and a kurtosis statistic of -0.406. Report has a skewness statistic of -0.412 and a kurtosis statistic of -0.522. Since the value for each of these statistics was within the recommended range of -2 to +2, there was no indication of any problems with univariate normality for either dependent variable. Further evidence of univariate normality was documented by the probability plots which showed close to straight lines (see Figure 4).

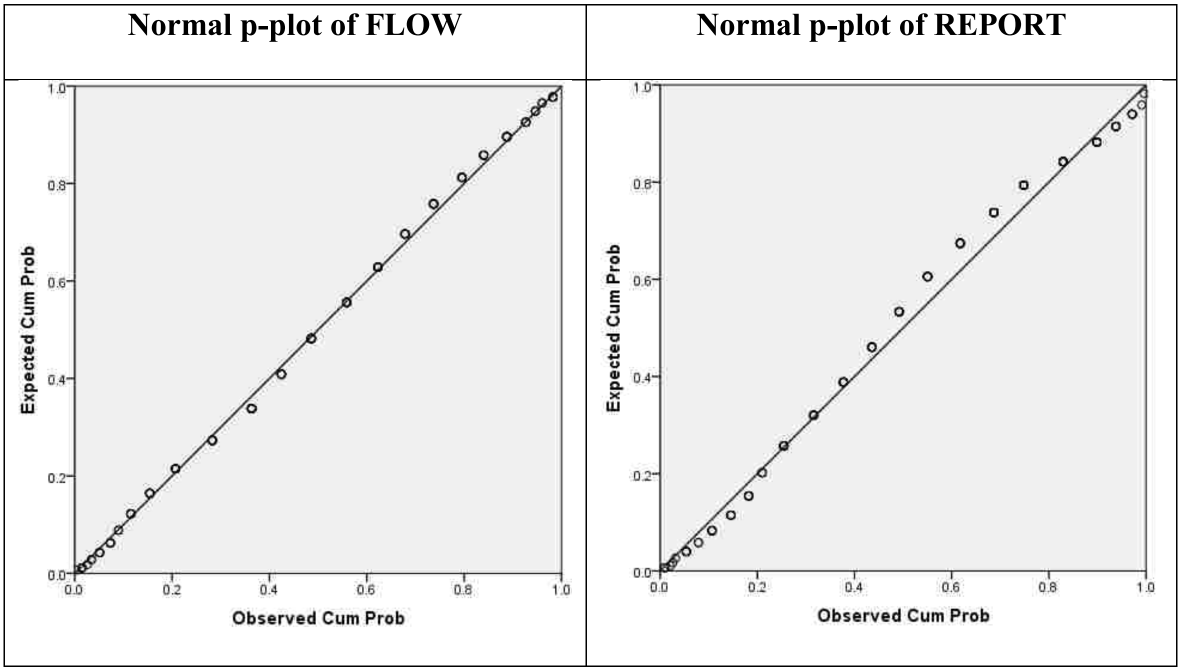


Figure 4: Normal probability plots for dependent variables.

To complete the multivariate normality check, a scatterplot was generated to investigate the bivariate normality of the pair of dependent variables. As shown in Figure 5, the elliptical shape represented bivariate normality.

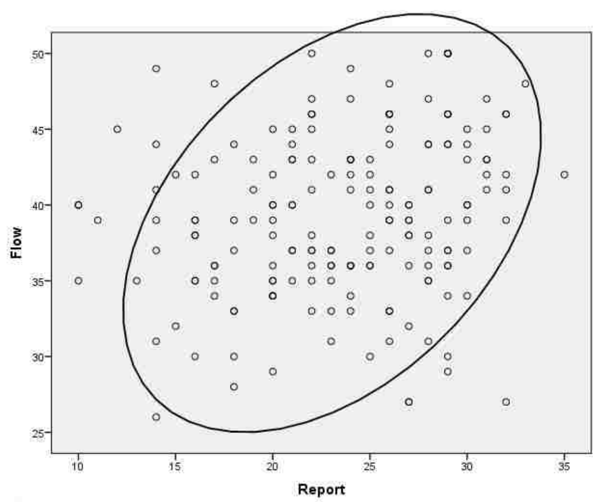


Figure 5: Scatterplot for dependent variables.

Next, the researcher tested for homogeneity of regression and the significance of regression using math PSSA score as the covariate. The interaction term of treatment group (independent variable) with math score (covariate) was not significant, (*Wilks'* $\lambda = .997$, $F(2, 174) = .227$, $p = .797$), so the homogeneity of regression assumption was met. Also, the test of significance of regression was significant (*Wilks'* $\lambda = .913$, $F(2, 175) = 8.346$, $p < .001$).

Lastly, the researcher tested for homogeneity of regression and the significance of regression using reading PSSA score as the covariate. The interaction term of treatment group (independent variable) with reading score (covariate) was not significant, (*Wilks'* $\lambda = .986$, $F(2, 174) = 1.218$, $p = .298$), so the homogeneity of regression assumption was met. Also, the test of significance of regression was significant (*Wilks'* $\lambda = .880$, $F(2, 175) = 11.982$, $p < .001$). With all assumptions met, the next step was testing for equality of adjusted means in the MANCOVA analyses.

RQ1: Multivariate significance between treatments. The first research question was: Do students in the INPLACE mobile game compared to the control activity report a more substantive flow-like experience and create culminating artifacts with more evidence of scientific practices? The primary hypothesis for this question was that the experimental group would show higher levels of flow experience and higher report scores; the MANCOVA tested for this as multivariate significance. Since classes were assigned to treatment groups intact and since the school district uses math PSSA scores to determine students' track level in school, math PSSA scores were used as the covariate to make group comparison fairer. Findings from the MANCOVA indicated that when controlling for math achievement, there was a statistically significant multivariate difference between treatment groups (*Wilks'* $\lambda = .961$, $F(2, 175) = 3.578$, $p = .03$).

The first follow-up question was: Are there any differences between treatment groups for flow experience or culminating artifact score when investigated separately? To answer question RQ1a, univariate tests were conducted to determine which variables (FLOW or REPORT) were contributing to the multivariate significance, results are shown in Table 6.

Table 6

Univariate ANOVA Tests for MANCOVA with Math Covariate

Measure	SS (contrast)	*df*	SS (error)	*F*	*p*	η^2
Flow experience	126.297	1	4987.080	4.457	.036	.025
Report score	103.148	1	4544.878	3.994	.047	.022

Follow-up univariate analyses indicated treatment group differences on both flow score (F(1, 176) = 4.457, p = .036) and artifact score (F(1, 176) = 3.994, p = .047); therefore, both dependent variables were contributing to the multivariate significance. When controlling for math achievement, the adjusted mean scores of flow for experimental participants (M = 40.50, SD = .72) were significantly higher than control participants (M = 38.62, SD = .50). All original and adjusted means are displayed in Table 7. Additionally, adjusted means of artifact scores for experimental participants (M = 24.68, SD = .69) were significantly higher than those of control participants (M = 22.98, SD = .47). In assessing effect sizes, Cohen (1977) stipulated how to interpret the partial eta squares for univariate ANOVA tests; he characterized .01 as small and .06 as medium. Thus, the effect sizes for the measures reported above are in the small to medium range. This is a reasonable effect size given the fairly short duration of the implementation as well as the comparison of an active, exploratory, collaborative lab activity to the same style of learning framed within an INPLACE mobile game; a large effect size was not anticipated.

Table 7

Descriptive Statistics with Adjusted Means for Dependent Variables

Variable	M (*n*)	SD	Adjusted means	Adjusted means
Flow experience score				
Experiment	40.51 (59)	6.62	40.50[a]	40.59[b]
Control	38.62 (120)	4.53	38.62[a]	38.58[b]
Total	39.24 (179)	5.37		
Report score				
Experiment	25.41 (59)	4.78	24.68[a]	24.80[b]
Control	22.63 (120)	5.53	22.98[a]	22.92[b]
Total	23.54 (179)	5.44		

[a] Adjusted means using math PSSA score as covariate.
[b] Adjusted means using reading PSSA score as covariate.

The second follow-up question was: Do reading achievement scores influence any of the above relationships? To investigate question RQ1b, a 1 by 2 factorial multivariate analysis of covariance (MANCOVA) was implemented and analyzed using SPSS. Overall, reading and writing ability may have influenced the relationships analyzed within the original MANCOVA, so to answer this research question, the original MANCOVA analysis was re-run with reading PSSA scores as the covariate instead of math PSSA scores. Adjusted means for this analysis are reported in Table 7. Findings from the MANCOVA indicated that when controlling for reading achievement, there was a statistically significant multivariate difference between treatment groups (*Wilks'* $\lambda = .952$, $F(2, 175) = 4.417$, $p = .013$).

Table 8

Univariate ANOVA Tests for MANCOVA with Reading Covariate

Measure	SS (contrast)	*df*	SS (error)	*F*	*p*	η^2
Flow experience	151.148	1	4977.477	5.344	.022	.029
Report score	132.060	1	4418.982	5.260	.023	.029

Univariate tests were conducted to determine which variables (FLOW or REPORT) were contributing to the multivariate significance; results are shown in Table 8. Follow-up univariate analyses indicated treatment group differences on both flow score ($F(1, 176) = 5.344$, $p = .022$) and artifact score ($F(1, 176) = 5.260$, $p = .023$); therefore, both dependent variables were contributing to the multivariate significance similar to the original MANCOVA analysis with math score as the covariate. When controlling for reading achievement, the adjusted mean scores of flow for experimental participants ($M = 40.59$, $SD = .71$) were significantly higher than control participants ($M = 38.58$, $SD = .49$). Additionally, adjusted means of artifact scores for experimental participants ($M = 24.80$, $SD = .67$) were significantly higher than those of control participants ($M = 22.92$, $SD = .46$). Similar to the effect sizes reported in the earlier analysis, effect sizes in this analysis are also in the small to medium range.

To check for any teacher effects in the data, intra-class correlations were output for teacher with report score, and teacher with flow score. Neither of the correlations was meaningful; therefore, students' report scores and flow scores were not related to which teacher they had for class. Additionally, intra-class correlations were output to ensure that the student team for which the student was assigned did not relate to that student's report score or flow score; the correlation between team membership and report score was not meaningful and neither was the correlation between team membership and flow score.

RQ2: Significance of flow as predictor of learning. The second research question was: Does flow experience predict culminating artifact score? The primary hypothesis for this question was that flow would predict some amount of variance in report scores for all students, while controlling for math achievement and treatment group. Descriptive statistics for the predictor variables and the outcome variable are included in Table 9.

Table 9

Descriptive Statistics for Predictors and Outcome Variable

Variable	M	SD	Observed Range	Possible Range
Math PSSA score	1555.78	259.42	1057.00-2308.00	727.00- 2548.00[a]
Flow experience	39.24	5.37	26.00-50.00	10.00-50.00
Report score	23.54	5.44	10.00-35.00	0.00-36.00

Note. Treatment group was excluded from this table because it is a dichotomous variable.
[a] The range was derived from the statewide minimum and maximum results.

A multiple hierarchical regression model was selected to test for this significance. Before beginning the regression analysis, assumptions were tested. First, data was checked for multicollinearity. Pearson correlations are reported in Table 10. Since all the values were low enough, all the predictors were kept in the model.

Table 10

Correlations between Predictors and Outcome Variable

Measure	1	2	3	4
1. Treatment group[a]	-	.314	.166	.241
2. Math PSSA score		-	.055	.344
3. Flow experience			-	.201
4. Report score				-

[a] Treatment group was dummy coded using control = 0 and experiment = 1.

Then, Cook's D values were checked; they were all below 1, so no outliers were biasing the data set. Finally, the data was checked for homoelasticity; the standardized plot revealed an even distribution.

Based on the MANCOVA results, reports scores were significantly different between treatment groups when controlling for standardized achievement in math and reading; therefore, treatment group was a possible confounding variable, so it was included in Block 1 of the hierarchical regression analysis along with math achievement. Flow experience was the variable of interest so it was added to the model in Block 2. Results are displayed in Table 11.

Table 11

Results for Multiple Hierarchical Regression Analysis Predicting Report Score (N = 179)

Predictor	$R^2\Delta$	$F\Delta$	*df*	$p\Delta$	*B*	*SE B*	*β*
Model 1	.138	14.087	2, 176	<.001			
Treatment group[a]					1.701	.851	.147*
Math PSSA score					.006	.002	.298***
Model 2	.164	5.511	1, 175	.020			
Treatment group[a]					1.387	.851	.120
Math PSSA score					.006	.002	.297***
Flow experience					.167	.071	.165*

Note. *B* = unstandardized regression coefficient; *SE* = standard error.
[a] Treatment group was dummy coded using control = 0 and experiment = 1.
*** $p < .001$
* $p < .05$

To investigate how well flow experience predicts report score, when controlling for treatment group and math achievement, a hierarchical linear regression was computed. In model 1, treatment group and math achievement were entered; they significantly predicted report score, $F(2, 176) = 14.09$, $p < .001$, adjusted $R^2 = .138$. By knowing the treatment group and math achievement, 13.8% of the variance in report score could be predicted. When flow experience was added, it significantly improved the prediction; R^2 change = .026, $F(1, 175) = 5.511$, $p = .02$. The semi-partial (part) correlation of flow experience was .162; therefore, flow experience uniquely predicted approximately 3% of the variance in report score. While math achievement was a significant predictor in both models, treatment group was a significant predicator in Model 1 but was no longer a significant predicator in Model 2. Overall, the entire set of variables—including math achievement, treatment group, and flow experience—significantly predicted report score, $F(3, 175) = 11.469$, $p < .001$, adjusted $R^2 = .164$.

Secondary Analysis

This section contains the within-case and cross-case analyses for the student teams that were examined as case studies. First, the within-case analysis for each team includes a case overview in summary form with quotations and photographic evidence followed by a data matrix. Second, the cross-case analysis represents all cases in a meta-data matrix. Finally, to answer the research questions, the findings from the cross-case synthesis are discussed and documented with additional summary displays and student quotations.

Description of Sample

For the qualitative analysis, the student team was the unit of analysis. Table 12 displays a demographic summary of all participating teams. Recall from Chapter 3 that the student teams were chosen using purposeful random sampling in order to represent a continuum of achievement levels in both treatment groups, yet maintain credibility of findings.

Table 12

Demographic Summary of Participant Teams

Case Study Identifier	Treatment (Achievement)	Gender Representation	Period (Teacher)	Math PSSA Range (Mean)	Reading PSSA Range (Mean)	Report Score Range (Mean)
Team C1	Control (Above Average)	Males = 2 Females = 2	Period 7 (Teacher A)	1499-1874 (1640)	1413-2091 (1678)	16-27 (21.75)
Team C2	Control (Average)	Males = 2 Females = 2	Period 1 (Teacher B)	1437-1739 (1591)	1413-1854 (1627)	27-29[a] (28.00)
Team C3	Control (Below Average)	Males = 2 Females = 1	Period 5 (Teacher B)	1191-1325 (1258)	1053-1092 (1073)	22-24 (22.67)
Team E1	Experiment (Above Average)	Males = 0 Females = 4	Period 5 (Teacher A)	1603-2148 (1898)	1691-1788 (1764)	27-32 (29.25)
Team E2	Experiment (Average)	Males = 3 Females = 1	Period 2 (Teacher B)	1339-1834 (1653)	1413-1788 (1606)	22-27 (24.50)
Team E3	Experiment (Below Average)	Males = 3 Females = 1	Period 1 (Teacher A)	1149-1259 (1187)	1032-1240 (1209)	21-22 (21.75)

[a] students were observed directly copying answers from each other in this team.

Within-Case Analysis

Case studies were conducted in order to investigate the communication responses and scientific practices that took place during the intervention. Communication responses that occurred in team conversations were categorized as accept, discuss, and reject. The use of this code structure built on the work of Barron (2003). When a student agreed with the speaker, supported the idea, or proposed a next step, then the interaction was coded as *accept*. When interactions facilitated further discussion, such as questioning an idea, asking for clarification, or challenging an idea with new information, then the interaction was coded as *discuss*. When a student rejected an idea or interacted in any way that would not facilitate further discussion, then the interaction was coded as *reject*. For examples of student dialogue coded with accept, discuss, and reject, please refer to Table 13.

Table 13

Coding Definitions for Communication Responses with Examples

Response	Definition	Examples
Accept	When a student agreed with the speaker, supported the idea, or proposed a next step	S1: He said there's no wrong answer. S2: Exactly. There's no wrong answer. (*Team C1 Conversation, 109:110*) S1: Okay, wait. Do we dip this in? S2: Yeah, when it's like liquid. (*Team C2 Conversation, 418:419*) S1: Number 7? S2: Yeah. (*Team C3 Conversation, 1109:1110*).
Discuss	When a student questioned an idea, asked for clarification, or challenged an idea with new information	S1: Yeah, it turned black and hardened. S2: So maybe cornstarch and sugar. (*Team C1 Conversation, 944:945*) S1: Why aren't we doing B? S2: Because he picked A. We can just do whichever one we want, but William picked A. (*Team C2 Conversation, 1589:1590*) S1: Where do we write that? S2: Middle of the box. (*Team C3 Conversation, 10:11*)
Reject	When a student rejected an idea or interacted in any way that would not facilitate further discussion	S1: It's fine. S2: No, it isn't. (*Team C1 Conversation, 183:185*) S1: Wait, can I see it real quick? S2: No, it's my turn. It's my turn. (*Team C2 Conversation, 903:905*) S1: It's shrinking. S2: It can't shrink unless it's melting. (*Team C3 Conversation, 627:628*)

Based on code reports, occurrences were categorized into levels of low (under 7), moderate-low (7-14), moderate (15-22), moderate-high (23-30), high (31-38), and very high (over 38) for each response type. Scientific practices that occurred in team conversations were coded to align directly to the scientific practices from the NRC (2012a). For each practice, the type of dialogue that would qualify as representing that practice was defined. For coding definitions of each practice along with examples of student dialogue, refer to Table 14.

Table 14

Coding Definitions for Scientific Practices with Examples

Number	Name	Definition	Example
Practice 1	Defining the Problem	When students tried to determine what needed to be answered or discussed what was known about the investigation	We will test three powders, do various experiments to ultimately conclude which powder is which. (*Team C1 Conversation, 41*) There was a theft at the cafeteria. (*Team E2 Conversation, 45*)
Practice 3	Planning the Investigation	When students discussed their investigation plan or what information they needed to record	Okay, using the first powder we can get, we will do the four tests - heat, pH, vinegar and iodine - and then keeping them in that order. (*Team C2 Conversation, 32*) Write the steps that you'll perform to identify the mystery powder. Well let's just say go talk to the janitor. (*Team E1 Conversation, 89*)
Practice 4	Interpreting Data	When students discussed characteristics of the experiments they were observing	It looks like boogers and snot. (*Team C2 Conversation, 405*) It's turning black. (*Team C3 Conversation, 352*) This is the baking soda one. So for that one it reacted and started to bubble. (*Team E3 Conversation, 802*) It turns yellow and then it turns liquidy and white again. (*Team E2 Conversation, 723*)
Practice 6	Constructing Explanations	When students constructed an explanation in order to explain the relationship between data	It was just bubbling like the baking soda. (Team C2 Conversation, 1149) Okay, well, this is fizzing. The only one that fizzed is baking soda. (*Team E1 Conversation, 581:582*)
Practice 7	Arguing with Evidence	When students supported or refuted an argument by citing relevant evidence	I think it's going to be sugar and baking soda, because it smelled…and it was just bubbling like the baking soda. (*Team C2 Conversation, 1149*) The cornstarch burned. So the only thing that burns is the cornstarch, so it has to be the cornstarch. (*Team E2 Conversation, 1351*)

Based on code reports, occurrences were categorized into levels of low (1-4), moderate (5-8), high (9-14), and very high (over 14) for each scientific practice. Case studies were also conducted in order to investigate additional differences between treatment groups during the intervention at the team level. The first significant difference that the researcher noticed was the general language style of the treatment groups. Students in the control frequently told each other what to do. They were using language such as, "don't reach across the table like that. Here—give it" (*Team C1 Conversation, 510*), "put the whole entire thing in" (*Team C2 Conversation, 122*), and "go get the other one" (*Team C3 Conversation, 285*). To capture this type of directive language, a new code was created called *commands*. Also, the researcher noticed that students in the game were addressing the team collectively, rather than a specific team member. They were also referring to the team as an entity with words such as "we," "we're," and "let's." To capture this language style, a new code was created called *communal language*.

The second difference was that students in both treatment groups were having off-topic conversations and managing some level of confusion. Students in both treatments had conversations about their clothes, after-school activities, television shows, movies, and popular music—to name a few. To capture these types of off-topic conversations, a new code was created called *off-topic*. Additionally, the researcher noticed that students were verbalizing their confusion with questions such as "Are we supposed to do the next one?" and confirming their uncertainty with statements such as "I have no idea" (*Team C2 Conversation, 19:22*). To capture these comments pertaining to confusion, a new code was created called *confusion*.

Lastly, teachers were involved with the treatment groups in extremely different ways. Control teams received a lot of direct instruction and small group scaffolding, as well as full class reviews. The game teams received very little—if any—of this type of teacher involvement.

The next six sections will summarize how each of these concepts was showcased for each case study team; each overview will conclude with a within-case display for that case.

Case study: Above average control team (C1). Math achievement was above average for this control team; the team consisted of two boys and two girls. They met during Period 7 with Teacher A. In general, one boy did not want to do anything, while the other one kept walking away from his team and disappearing. One of the girls was very talkative with other people, while one girl was generally on-task however she seemed to have a slight attitude at times.

During the activity, no leader emerged. Teammates somewhat disagreed with each other since they had a moderate level of reject responses. They did not insult each other or shows signs of resentment, although they did have a moderate level of commands such as “hey, don’t put it on there yet” (*Team C1 Conversation, 509*) and “keep burning it” (*Team C1 Conversation, 633*).

Overall, they were fairly democratic. They had a moderate level of discuss responses. Refer to Figure 6 for evidence of the group working together in two different scenarios. While their interactions were democratic, they were fairly ineffective. They only had a moderate level of accept responses, so they supported each other’s ideas about half of the time while rejecting them the other half of the time. They were not functioning as a completely cohesive group shown by their moderate level of communal language. For example, the team had difficulty keeping their powders organized. Their confusion could have been ameliorated if someone had control of the situation. Yet, one member’s mistake would ultimately become a mistake for the team.

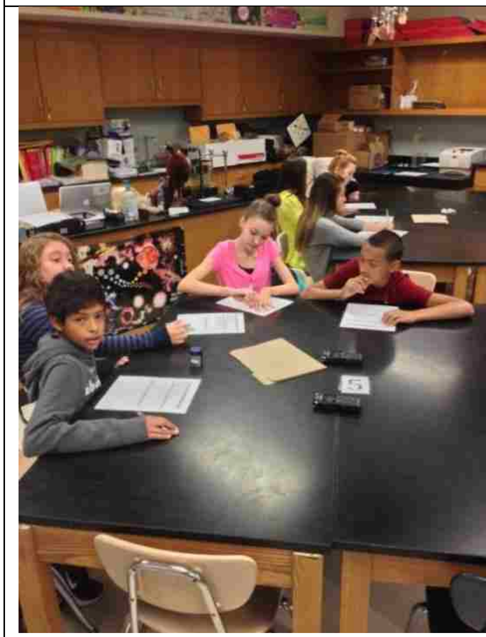
Team during planning	Team while experimenting
	

Figure 6: Team C1 under different working conditions.

The biggest problem for this team was their confusion. Their achievement level was above average so they were able to follow the instructions sheet, gather the supplies, and conduct the experiments. They had a very-high level of discussing their observations during the experiments. However, when it was time to synthesize the information, they had a moderate to high level of confusion represented in the discussion. They would get confused and talk off-topic. Then, they would realize how far behind they were and rush at the end of the period. Sometimes, if a team member was not at the table at that time, they would be left behind. Generally, when they were confused, they would go ahead anyway, and their confusion compounded. When they did ask the teacher for assistance, they did not incorporate teacher comments into their reports. For example, the team confirmed with the teacher that their mystery powder had two powders, yet the team wrote up their final conclusions stating that they were three powders in the mystery powder.

The team had a moderate level of off-topic comments in their discussion. However, their off-topic behavior was not necessarily captured in their conversations. There was a lot of movement between tables and there was doodling on the back of their report papers. Refer to Figure 7 for photographic evidence of their off-task doodling.

Figure 7: Evidence of doodling in Team C1.

As far as scientific practice, this team had a very high level of discussing their data (Practice #4); they followed the instructions and got the experiments done although they had trouble staying organized and keeping their powders straight. Yet, one scientific practice without the others is not effective. They only had a moderate level of defining the problem (Practice #1), constructing explanations (Practice #6), and trying to argue with evidence (Practice #7) but having been so confused for so long they could not overcome their confusion and synthesize the information effectively. Although the team had a high level of planning their investigation (Practice #3), the discussion lacked substance and specific details. Their reports from this team showed a lack of understanding of the main goal of the experiment as well as most of the details. Thus, it seems the team needed a better understanding of the overall direction of the experiment in order to alleviate their confusion.

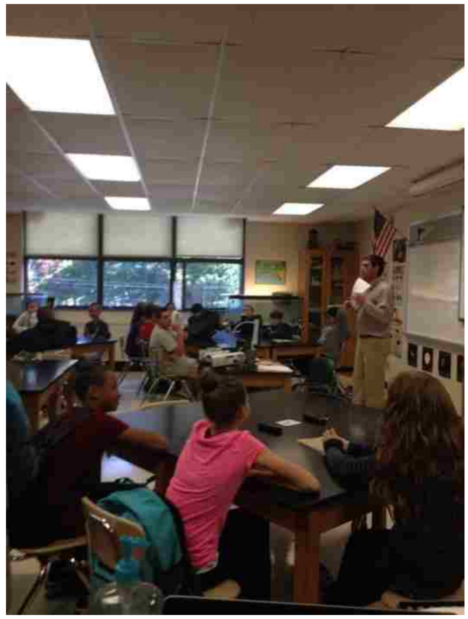


Figure 8: Teacher A giving whole-class instruction.

Unfortunately, the teacher was minimally involved with this team; therefore, teacher involvement could not alleviate confusion within the group. Sometimes, at the beginning of the period, the teacher would offer some brief instructions to the whole class (see Figure 8). He would start class assuming that everyone was paying attention, even though this was not always the case. Once his lost control early in the period, he would struggle for the rest of the period to gain it back. In general, Teacher A did not implement enough classroom management strategies. Ultimately, it seemed like the students knew that they could take advantage of his lack of experience. For a within-case summary display of Team C1, see Table 15.

Table 15

Case Study Overview for Team C1

Communication Responses	Scientific Practices	Group Dynamics	Nature of Non-Scientific Comments	Teacher Involvement
Moderate reject Moderate discuss Moderate accept	Interpreting data: Very high Planning investigation: High Arguing with evidence: Moderate Constructing explanations: Moderate Defining the problem: Moderate	No leader emerged. Democratic but not effective.	Moderate-high confusion Moderate off-topic Moderate commands Moderate communal	Some instruction Some reviewing Assistance as needed

Case study: Average control team (C2). Math achievement was average for this control team; the team consisted of two boys and two girls. They met during Period 1 with Teacher B. In the beginning, the girls were somewhat hesitant to talk. One boy wanted to take leadership and did not want anyone else to do anything; the other boy seemed willing to defer to the leader boy.

Over the course of the activity, the strong-willed boy controlled the leadership; he was a very controlling, demanding leader and an ineffective communicator. Team members disagreed with each other often and did not support each other's ideas; they had a high level of reject responses. They also insulted each other frequently and showed signs of resentment such as "Oh my God. You are so stupid." (*Team C2 Conversation, 551*) and "Accidents happen, like [boy's name]. You were an accident." (*Team C2 Conversation, 1082*). They continued throughout the activity to direct each other with a high level of commands such as "you don't have it close enough…put it down closer" (*Team C2 Conversation, 558*) and "hey, I told you to do it…put it on the like top" (*Team C2 Conversation, 439*).

A lot of the group difficulty seemed to stem from fighting over roles and responsibilities. Here is a sample conversation about who will conduct the heat test:

S1: I want to do the heat test.
S2: I called it yesterday.
S1: I called it today.
S3: Are you guys fighting for which?
S2: I want to do the heat test.
S1: Okay, that's it.
(*Team C2 Conversation, 242:247*)

Overall, this was a dictatorship with the leader boy in the control position. The team had a moderate level of discuss responses with a high level of reject responses. They also had a moderate to low level of accept responses. In brief, the discussions they did have were not productive. They were clearly not functioning well as a group showing a low level of communal language. Over the course of the week, contention built up until the one girl stood up for herself. She and the leader boy exchanged these words when she was clearly upset at him:

> Girl: Why are you doing everything?
> Boy: Because no one else is getting up and doing anything.
> Girl: Because you're not letting us.
> (*Team C2 Conversation, 1172:1174*)

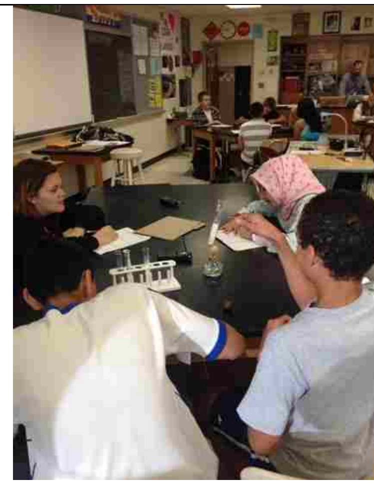
Day 1	**Day 2**	**Day 3**
		

Figure 9: Team C2 leader conducted most of the experiments himself.

The biggest problems for this team were their group dynamics and in-team communication. The leader boy wanted to do everything at all times, as shown in Figure 9. The few times he allowed others to conduct experiments he was overbearing and easily aggravated. Their inability to function well as a team led to a moderate to high level of confusion for certain team members. For example, when testing the mystery power, the leader boy rushed through all the experiments while the others paid partial attention but mostly just reprimanded him for going too fast. When the team tried to synthesize the information, the acquiescent boy was confused and continued to ask his teammates for assistance. One girl was apt to try to explain the answer while the leader boy was more apt to say: "Just read my paper" (*Team C2 Conversation, 328*). All reports were almost identical and copying was observed. The team had a moderate to high level of off-topic comments in their discussion, but this did not seem to detract from their learning as much as the group dynamics issue.

Confusion in this team was the result of a several issues. Similar to other control teams, they had trouble keeping their known powders organized. They also did not necessarily write down their observations at the time they conducted the experiment. After completing several experiments, the team would discuss which powder revealed which results, since they had not written down the information immediately. Ultimately, there was a lot of confusion over what they needed to do, who would be responsible for doing it, which powders they were actually testing, and what they were going to write down.

As far as scientific practice, this team had a low level of describing the experiment (Practice #1) yet a high level of planning their investigation (Practice #3). The teacher allocated time for both these discussions; he even intervened during the team discussion to ask about their ideas. Since most of the class periods were spent conducting experiments, this team had a very high level of discussing their observations of the experiments (Practice #4), similar to case study C1. When the team made observations, they were valid scientific observations. At the very end of the activity, the team was able to discuss some observations of the mystery powder and connect it to previous observations of sugar and baking soda. There was a moderate level of constructing explanations (Practice #6). There was only a low level of arguing with evidence (Practice #7). Student reports—however—showed a higher level of Practice #7 since this team retrieved some of their insights by watching and listening to other teams.

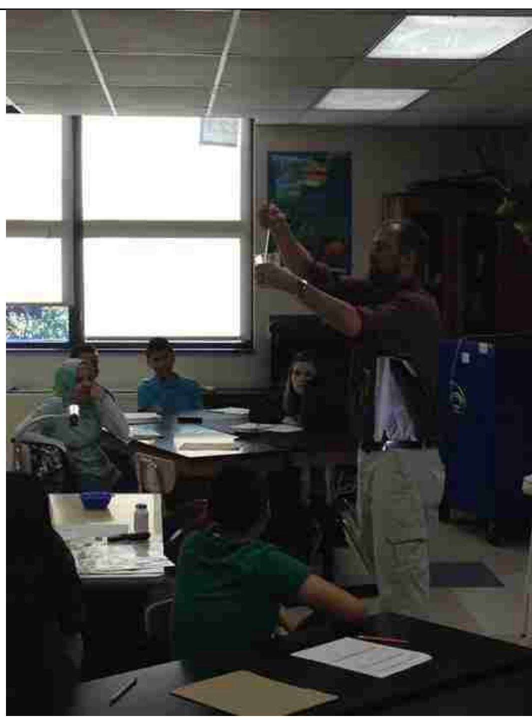
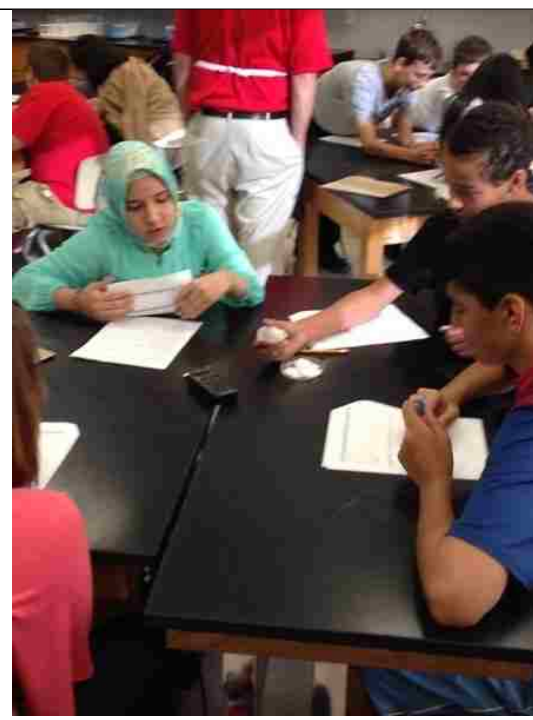
General overview	**Group review**
	
Group demonstration	**Assisting as needed**
	

Figure 10: Different teaching techniques exhibited by Teacher B.

Finally, this team had the benefit of the more veteran teacher. Teacher B implemented several classroom management techniques and mostly maintained control over his classes. For visual evidence of the variety of teaching techniques, refer to Figure 10. In the beginning of the week, Teacher B began with a content and instructional overview. Throughout the class periods, the teacher would circulate to teams and have guiding conversation with them to make sure that they were building an understanding of the activity. If time permitted at the end of the period, the teacher would reconvene the class and review what was learned. Ultimately, it seemed that if the students paid attention to the teacher, they could glean a lot of helpful information and clarify their understanding. Refer to Table 16 below for a within-case summary display of Team C2.

Table 16

Case Study Overview for Team C2

Communication Responses	Scientific Practices	Group Dynamics	Nature of Non-Scientific Comments	Teacher Involvement
High reject Moderate discuss Moderate-low accept	Interpreting data: Very high Planning investigation: High Constructing explanations: Moderate Defining the problem: Low Arguing with evidence: Low	One boy took leadership. Ineffective communicator and very controlling. Teammates showed resentment.	High commands Moderate-high confusion Moderate-high off-topic Low communal	Direct instruction; scaffolding; reviewing

Case study: Below average control team (C3). Math achievement was below average for this control team; the team consisted of two boys and one girl. They met during Period 5 with Teacher B. In the beginning, the boys were kind of quiet. The girl seemed knowledgeable and interested in science and took a leadership role. She would delegate the boys to go get supplies. One boy liked to get up and talk to other tables. The girl acted as the team motivator; she was the reason the team kept moving forward.

Over the course of the activity, the girl remained the leader; she was a decent communicator and delegator, not controlling like the leader of C2. Team members disagreed with each other fairly often; they had a moderate to high level of reject responses. Yet, they also had a moderate to high level of discuss responses as well as a moderate level of accept responses. So, there was a lot of discussion but a mixed level of support for each other's ideas. As far as team communication, they did not insult each other too much or show signs of resentment. However, there was not a lot of group cohesion since the communal language was low.

Since the girl had the most understanding of how to accomplish the activity, she directed her teammates with a high level of commands such as "here, do the last one" (*Team C3 Conversation, 360*) and "you have to hold it from the bottom" (*Team C3 Conversation, 842*). Her commands were often directing the boys on how to execute the hands-on activities. Sometimes the leader girl led the experiment; at other times, she directed and delegated her teammates to do so, as shown in Figure 11. Thorough the activity, there was relatively little contention over roles and responsibilities. This team was able to negotiate responsibilities with each other, for example: "Let me do the iodine. Will you give me that stuff?" (*Team C3 Conversation, 1076*).

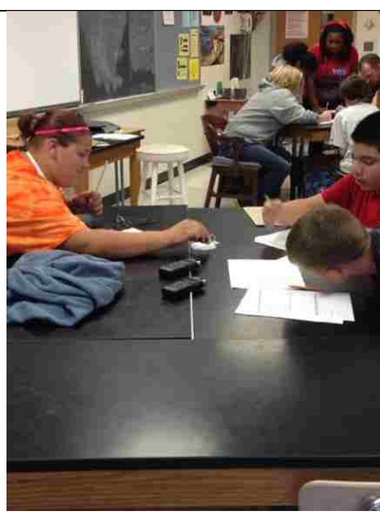
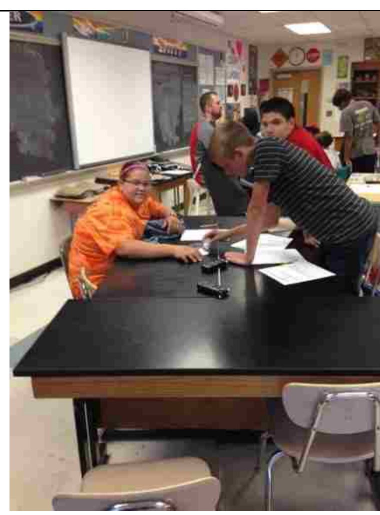
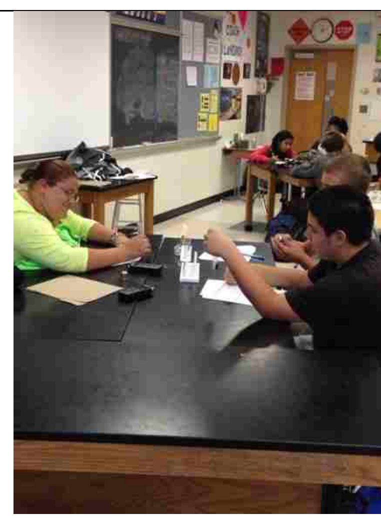
Girl leading	**Boy 1 leading**	**Boy 2 leading**
		

Figure 11: Team C3 girl leader encouraged even delegation of responsibilities.

The biggest problems for this team were their high levels of off-topic conversations and moderate confusion. When the leader girl left the table, the boys would talk off-topic. At times, the girl leader tried to steer the team back on track: “We’re not talking about that. We’re talking about science” (*Team C3 Conversation, 104*). But at other times, the boys were so entrenched in their off-topic conversations that the girl was not able to get them back on-task. For example, during the second half of third class period, the conversation was supposed to be about summarizing the data table and writing hypotheses; instead, it went completely off-topic showcasing no scientific practice. This low level of effort and understanding was evident in their reports.

As far as general scientific practice thorough out the activity, this team had a low level of describing the experiment (Practice #1), constructing explanations (Practice #6), and arguing with evidence (Practice #7). The few occurrences of Practice #6 were all spoken by the leader girl and the one occurrence of Practice #7 was actually facilitated by the teacher:

> T: What’s the mystery powder?
> S1: Cornstarch and baking soda.
> T: Okay, why?
> S1: Because all the test results and reactions have something to do with the cornstarch and baking.
> T: Be specific. Why? What results?
> S2: Because the heat test, the cornstarch and the baking soda turned black.
> T: Okay. What else?
> S1: And the vinegar test, the baking soda sort of fizzled, and the iodine test the cornstarch turned black.
> T: There you go. That’s specific. That’s what you should be writing.
> (*Team C3 Conversation, 1228:1237*)

So, Practice #6 and Practice #7 were almost none existent without teacher assistance. Team C3 did reveal a high level of planning their investigation (Practice #3) though. The teacher allocated plenty of time for this discussion; he even asked the team about their ideas and gave them positive feedback on their plan. Similar to the other control case studies, most of the class periods were spent conducting experiments, so this team had a very-high level of discussing their observations of the experiments (Practice #4). Unfortunately, when the leader girl spoke using scientific terms, the boys often protested about her use of "big words" (*Team C3 conversation, 135*); one boy even complained that he had a "little brain" (*Team C3 conversation, 137*). Similar to C2, this team retrieved some of their answers by talking to other teams. Yet, since there were two different mystery powder samples, the boy who thought he got the answers from another team, wrote up his final conclusion incorrectly because his team had a different mystery powder.

Finally, similar to Team C2, this team had the benefit of the more veteran teacher. As mentioned earlier, Teacher B implemented several classroom management techniques and mostly maintained control over his classes. This team probably had the most teacher involvement of any team. Every period, Teacher B began with an instructional overview, and when possible, reconvened the class at the end to review what was learned. At times, this even included essentially giving the answers, for example:

> T: If I add vinegar to the mystery powder and it fizzes, then what is in the mystery powder?
> S1: Sugar.
> S2: Baking soda.
> T: [I] just said baking soda.
> (*Team C3 Conversation, 585:588*)

Throughout the class periods, the teacher circulated to teams to have guided conversations with team C3; this interaction was essential for keeping this team on-task ("After you put down, you put the vinegar on powder, write down what you observe." *Team C3 Conversation, 318*) and promoting understanding of the activity. Refer to Table 17 below for a within-case summary display of Team C3.

Table 17

Case Study Overview for Team C3

Communication Responses	Scientific Practices	Group Dynamics	Nature of Non-Scientific Comments	Teacher Involvement
Moderate-high reject Moderate-high discuss Moderate accept	Interpreting data: Very high Planning investigation: High Constructing explanations: Low Defining the problem: Low Arguing with evidence: Low	Girl leader emerged. Decent communicator and delegator. Boys followed willingly while at times resisting.	High commands High off-topic Moderate confusion Low communal	Direct instruction; scaffolding; reviewing

Case study: Above average experiment team (E1). This team in the experiment consisted of four girls; they had above average math achievement. They met during Period 5 with Teacher A. In the beginning, the girls seemed very excited about doing the activity with comments such as "It seems so fun. I'm so excited" (*Team E1 Conversation, 35*) and "this is so cool" (*Team E1 Conversation, 44*). The girls were generally on-task and seemed to stay together and work well together as a team.

Over the course of the activity, no leader emerged. Team members did not generally disagree with each other; they had a low level of reject responses. They also did not insult each other or shows signs of resentment. In fact, they had a high level of discuss responses and a high level of accept responses showing that they discussed ideas collectively and supported each other's thinking.

Overall, this all-girl team was very communal. They were physically communal sitting or standing in close proximity to each other, as shown in Figure 12. They also had a high level of communal language and a low level of commands in their dialogue. Their communal language served not only to promote group cohesion but also to keep them on track towards their goals: "Okay, now we have to figure out what color" (*Team E1 Conversation, 424*). As shown in Figure 12, they even stayed together while waiting for their teammate to finish writing before moving on to the next game location.

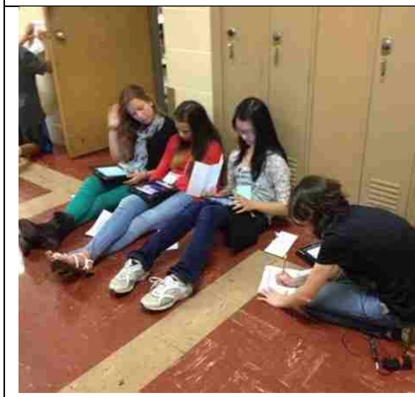
Working together	Waiting for teammate	Teacher assistance
		

Figure 12: Team E1 group interactions.

The girls were generally on-task for the duration of the activity. The girls revealed a moderate to low level of off-topic comments. They needed no direct instruction to stay on task. The teacher was only involved in their learning process when a question arose within the group; the team briefly used the teacher as a guide on the side as shown in Figure 12. The girls did not even need the teacher to ameliorate their moderate to low level of confusion; instead, they often worked through confusion through group discussion.

As far as scientific practice, this team had very high levels of discussing their data (Practice #4). While their overall number of occurrences may not have been as high as other teams, team E1 actually used Practice #4 in conjunction with other practices, specifically constructing explanation (Practice #6). For example, “Okay, well, this is fizzing. The only one that fizzed is baking soda” (*Team E1 Conversation, 582*). For 6 of the 21 occurrences of Practice #4, Practice #6 followed. In fact, team E1 was one of only two cases demonstrating a high level of Practice #6, constructing explanations. Team E1 only had moderate levels of describing their investigation (Practice #1) and trying to argue with evidence (Practice #7). Interestingly, E1 and E3 were the only teams that actually showcased these practices together. In both cases, while team members were discussing who the suspects were, they argued with evidence as they gave reasons for who might be suspect. Team E1 exhibited a low level of planning their investigation (Practice #3); this was the only practice that ranked a low level rating.

Overall, there was no primary problem for this team. In contrast to the C1 team that had the same teacher, the E1 team paid attention to the activity and excelled in their collaborative learning process; their report scores are the highest of any case study team. Unlike in C1 who struggled with persistent confusion based on a lack of understanding of the direction of the activity, E1 suffered little confusion and stayed on task towards their goal with the same teacher but within the game framework. In the C1 case study, the team showed that one scientific practice without the others is not effective; whereas in the E1 case study, the team revealed that moderate occurrences with integrated practices was extremely effective. Refer to Table 18 below for a within-case summary display of Team E1.

Table 18

Case Study Overview for Team E1

Communication Responses	Scientific Practices	Group Dynamics	Nature of Non-Scientific Comments	Teacher Involvement
High discuss High accept Low reject	Interpreting data: Very high Constructing explanations: High Arguing with evidence: Moderate Defining the problem: Moderate Planning investigation: Low	Group discussed and selected game roles. Group made decisions together.	High communal Moderate-low confusion Moderate-low off-topic Low commands	Very low teacher involvement. No direct instruction, scaffolding, or review. Only assistance as needed.

Case study: Average experiment team (E2). This team in the experiment consisted of three boys and one girl; they had average math achievement. They met during Period 2 with Teacher B. In general, one boy did not seem to get along entirely well with the team, although it in unclear why that was the case. In one instance, he took a partner and separated from the others and this angered his teammates: “They’re bad people. They left us. They split up. That’s illegal.” (*Team E2 Conversation, 463*).

Over the course of the activity, no leader emerged. Teammates disagreed with each other about half of the time since they had a moderate level of reject responses. Most reject responses were directed at the one boy who did not assimilate well with the group. The team did support each other's ideas more than half the time with a moderate to high level of accept responses. Their communication responses were primarily discussion based revealing a very high level of discuss responses. Sometimes they insulted each other showing a few signs of resentment with comments such as "No we don't, stupid" (*Team E2 Conversation, 53*) and "Come on, [boy's name]. You're such a girl" (*Team E2 Conversation, 1321*).

Working together in the hall	**Working together in class**
	

Figure 13: Team E2 group interactions.

Overall, their process of interaction was democratic and generally effective. To see pictures of this team interacting, refer to Figure 13. The team showcased a moderate level of commands; however, they seemed to be evenly distributed to a variety of members: "[name of student 1], come here. Come here" (*Team E2 Conversation, 82*); "[name of student 2], get back here" (*Team E2 Conversation, 298*); and "[name of student 3], stop that" (*Team E2*

Conversation, 307). Additionally, many of their commands occurred in the classroom when the team clearly struggled the most working together as a group. During experimentation, team members said commands such as “three…just put three in” (*Team E2 Conversation, 1173*) and “don’t put it all the way down” (*Team E2 Conversation, 1309*). Ultimately, their discuss and accept responses were high enough to overcome their moderate level of reject responses, commands, and insults. They actually functioned well as a cohesive group showing a high level of communal language. For example:

> S1: All right, so we have to go to the art room and the shop.
> S2: All right, guys, let’s go.
> (*Team E2 Conversation, 708:709*)

The biggest problem for this team was their moderate level of confusion. Although they did have a moderate level of off-topic discussions, they occurred during walking excursions to get to the next game code, so they did not detract from the learning process. The moderate confusion was present throughout the activity. Students in this team split up temporarily during the second day of gameplay; once they regrouped, it took a while to ascertain what information they had and still needed. Some early confusion also resulted from uncertainty about how to use the technology. Some teacher assistance was required to ameliorate confusion in this team.

As far as scientific practice, this team had very high levels of discussing their data (Practice #4). When they discussed their data, it was often a discussion with all members contributing what they knew. Discussing their data collectively also led to constructing explanations (Practice #6) and arguing using evidence (Practice #7); this team showed moderate levels for both these practices. This example demonstrates Practices #6 and #7:

> S1: Did the baking soda fizz a lot more than the cornstarch?
> S2: Here, let me check. So baking soda. Yeah, baking soda fizzed a lot more than; it fizzed a lot more than a lot of them
> ...

S2: But the cornstarch also did fizz.
S3: But it didn't turn black.
S2: Well with vinegar I meant, I'm talking about.
S3: So and the cornstarch burned. So the only thing that burns is the cornstarch, so it has to be the cornstarch.
…
S3: It burned. It's the only thing that burned. It has to be.
(*Team E2 Conversation, 1343:1356*)

Team E2 only had moderate levels of describing their investigation (Practice #1) and a low level of planning their investigation (Practice #3). Interestingly, this team exhibited the ability to synthesize their information well but did not spend a lot of time discussing information that was purely informational reporting. They must not have needed lengthy discussions about the incident in order to understand it because they all scored full marks on their reports for their descriptions.

Overall, this team struggled somewhat with group dynamics in situations that were outside of the game framework, such as conducting the mystery powder experiments. However, when it came to synthesizing the information and drawing conclusions collectively as a team, they excelled. Refer to Table 19 for a within-case summary display of Team E2.

Table 19

Case Study Overview for Team E2

Communication Responses	Scientific Practices	Group Dynamics	Nature of Non-Scientific Comments	Teacher Involvement
Very high discuss Moderate-high accept Moderate reject	Interpreting data: Very high Constructing explanations: High Arguing with evidence: Moderate Defining the problem: Moderate Planning investigation: Low	Group selected game roles. Not too many group issues during gameplay. Some issues in classroom.	High communal Moderate commands Moderate confusion Moderate off-topic	In the halls, teacher assistance with technology and corralling students at end of periods. In class, some instruction and scaffolding. No reviewing.

Case study: Below average experiment team (E3). This team in the experiment consisted of three boys and one girl; they had below average math achievment. They met during

Period 1 with Teacher A. In general, all the individuals in this team seemed to be quiet and reserved. It was also a group with some IEPs and therefore they had an aide assist them on the first day. However, one boy took the leadership role and taught the rest of his team how to use the technology and helped them to understand the content and fill out their worksheets.

Over the course of the activity, the one boy emerged as the leader. Team members very infrequently disagreed with each other with a low level of reject responses. However, there was not a lot of evidence for supporting each other's ideas either with a moderate to low level of accept responses. Their communication responses were primarily discussion based revealing a very high level of discuss responses. They rarely insulted each other and showed no signs of resentment. This is notable because in other teams where a leader emerged, there were higher levels of reject responses, insults, or resentment.

In the classroom	**At the table**
	

Figure 14: Team E3 in class.

Overall, their process of interaction was a blend of directed leadership and communal effort. To see pictures of this team, refer to Figure 14. The team exhibited a high level of communal language asking questions such as "where do we go now?"(*Team E3 Conversation,* *368*) and using phrases such as "now we have to tell him about the evidence" (*Team E3 Conversation, 573*). The team showcased a moderate level of commands; however, these commands were mostly part of the leader's instructional technique for helping his teammates. Here is a sample conversation where the leader tries to establish what it known about the powders and tests while guiding another teammate on where to look for information on his iPad:

> Leader: All right, so what happened with the vinegar? The vinegar, it dissolved a little bit. So put it dissolved in the box under sugar collected and under the vinegar column. It dissolved, that's it. All right. You have the iodine test.
> Teammate: What?
> Leader: Go into your inventory and look for the iodine test. Is that the sugar?
> Teammate: It's iodine test.
> Leader: No, that's the cornstarch. You have to go to the sugar. You have to go to the sugar slash iodine test. All right. So for that one it dissolved and it turned green. Now who has the heat test?
> (*Team E3 Conversation, 787:792*)

Additionally, this team had a low level of confusion and low level of off-topic comments. Of all the case study teams in the experiment, this team had the most teacher involvement for two reasons: (1) the teacher gave more directed instruction to this class period and (2) this team had a dedicated aide some days. Therefore, their confusion was low. When issues arose, the leader boy assisted his teammates or deferred to the expertise of the classroom teacher. Off-topic conversations were minimal and certainly did not detract from the learning process

As far as scientific practice, this team had very high levels of discussing their data (Practice #4) similar to all other case study teams. Rather than discussing their data as they collected it around the school, this team discussed the data when they arrived back at the classroom and were seated at their table. This team had the highest level of describing their investigation (Practice #1); they were the only team that showcased a high level of this practice. It was during their discussion of the incident when their conversation seemed to include all members participating evenly, rather than the leader steering the conversation. It was also this conversation that led to the team's first occurrence of Practice #7, arguing with evidence; one team member tried to provide evidence for why the janitor should be considered a suspect by stating "because when we were talking with him, he was …. He started …. because he was sick, and he might have …. because he's the only one in the building after 5:00" (*Team E3 Conversation, 463*). This team exhibited a moderate level of Practice # 7 overall, which aligned to the same level of the other experimental case studies. Team E3 only had low levels of planning their investigation (Practice #3) and constructing explanations (Practice #6).

The biggest problem for this team may have been the reserved nature of the majority of their members. There were not a lot of conversations where the team collectively generated ideas. In summation, this team seemed to achieve a group dynamics structure that worked well in terms of communal cohesion; they had low conflict and low confusion. However, the dynamics did not yield the most productive conversations since not all members were actively involved. This meant the team struggled somewhat to achieve a high level of synthesis at the end of the activity. Fortunately, the game narrative game the team a tangible means for showcasing a high level of Practice #1 and moderate level of Practice #7; these levels were notably higher than those of the comparable control team, C3. Refer to Table 20 below for a within-case summary display of Team E3.

Table 20

Case Study Overview for Team E3

Communication Responses	Scientific Practices	Group Dynamics	Nature of Non-Scientific Comments	Teacher Involvement
Very high discuss Moderate-low accept Low reject	Interpreting data: Very high Defining the problem: High Arguing with evidence: Moderate Constructing explanations: Low Planning investigation: Low	Everyone chose game roles; however, boy leader emerged…very good communicator and teacher. No resentment towards leader.	High communal Moderate commands Low confusion Low off-topic	Teacher assistance with technology. Scaffolding. Basic instruction. Brief reviewing. [NOTE: This group had a dedicated aide on the first day.]

Cross-Case Analysis

The previous section provided rich, thick descriptions of each team's scientific practices, communications responses, and leadership organization as well as teacher involvement in their learning process. The purpose of this section is to present a cross-case synthesis of the data from the individual cases in order to answer the following three research questions:

RQ3. How do in-team communication responses during the INPLACE mobile game compare to in-team responses during the control activity?

RQ4: How do scientific practices of teams in the INPLACE mobile game compare to those of teams in the control activity?

RQ5: How else are treatment groups different when qualitatively analyzed at the team level?

To begin the analysis, a case-ordered descriptive matrix (refer to Table 21) is provided with the control cases placed in the top three rows and ordered from above average to below average achievement level, and the experimental cases placed in the bottom three rows also ordered in descending achievement level. For each research question, a synthesis is provided for control and quasi-experimental cases supported by evidence from case studies and the meta-data matrix. Additional detail summaries are provided when needed. The final summation is presented last as the finding for the research question.

RQ3: Communication responses. The third research question is: How do in-team communication responses during the INPLACE mobile game compare to in-team responses during the control activity? In order to answer this question, a data analysis of the communication responses among teams within both treatment groups will be presented. First, this section will summarize the patterns of communication responses observed in the control group followed by the response patterns from the teams in the experiment. Lastly, a cross-case synthesis will be presented as the final finding for the research question.

Table 21

Cross-Case Analysis Summary

Team	Communication Responses	Scientific Practices	Group Dynamics	Nature of Non-Scientific Comments	Teacher Involvement
Control (Above Average)	Moderate discuss Moderate accept Moderate reject	Interpreting data: Very high Planning investigation: High Arguing with evidence: Moderate Constructing explanations: Moderate Defining the problem: Moderate	No leader emerged. Democratic but not effective.	Moderate communal Moderate commands Mod-high confusion Moderate off-topic	Some instruction; some reviewing; assistance as needed
Control (Average)	Moderate discuss Mod-low accept High reject	Interpreting data: Very high Planning investigation: High Constructing explanations: Moderate Defining the problem: Low Arguing with evidence: Low	One boy took leadership. Ineffective communicator and very controlling. Teammates showed resentment.	Low communal High commands Mod-high confusion Mod-high off-topic	Direct instruction; scaffolding; reviewing
Control (Below Average)	Mod-high discuss Moderate accept Mod-high reject	Interpreting data: Very high Planning investigation: High Constructing explanations: Low Defining the problem: Low Arguing with evidence: Low	Girl leader emerged. Decent communicator and delegator. Boys followed willingly while at times resisting.	Low communal High commands Moderate confusion High off-topic	Direct instruction; scaffolding; reviewing
Experiment (Above Average)	High discuss High accept Low reject	Interpreting data: Very high Constructing explanations: High Arguing with evidence: Moderate Defining the problem: Moderate Planning investigation: Low	Group discussed and selected game roles. Group made decisions together.	High communal Low commands Mod-low confusion Mod-low off-topic	Very low teacher involvement. No direct instruction, review, or scaffolding. Only assistance as needed.
Experiment (Average)	Very high discuss Mod-high accept Moderate reject	Interpreting data: Very high Constructing explanations: High Arguing with evidence: Moderate Defining the problem: Moderate Planning investigation: Low	Group selected game roles. Not too many group issues during gameplay. Some issues in classroom.	High communal Moderate commands Moderate confusion Moderate off-topic	No reviewing. In the halls, teacher assistance with technology and corralling students at end of periods. In class, some instruction and scaffolding.
Experiment (Below Average)	Very high discuss Mod-low accept Low reject	Interpreting data: Very high Defining the problem: High Arguing with evidence: Moderate Constructing explanations: Low Planning investigation: Low	Everyone chose game roles; however, boy leader emerged…very good communicator and teacher. No resentment towards leader.	High communal Moderate commands Low confusion Low off-topic	Teacher assistance with technology. Basic instruction; brief reviewing; scaffolding. [NOTE: This group had a dedicated aide on the first day.]

As described earlier, the communication responses that occurred in team conversations were categorized as accept, discuss, and reject. The use of this code structure built on the work of Barron (2003). When a student agreed with the speaker, supported the idea, or proposed a next step, then the interaction was coded as *accept*. When interactions facilitated further discussion, such as questioning an idea, asking for clarification, or challenging an idea with new information, then the interaction was coded as *discuss*. When a student rejected an idea or interacted in any way that would not facilitate further discussion, then the interaction was coded as *reject*. Based on code reports, teams were categorized into levels of low, moderate-low, moderate, moderate-high, high, and very high for each response type. For a summary of all the occurrences of these types of interactions within each team, please refer to Table 22.

Table 22

Content-Analytic Summary Table of Communication Responses for All Cases

	Above Average	Average	Below Average
Control	Moderate reject (16) Moderate discuss (21) Moderate accept (15)	High reject (37) Moderate discuss (19) Moderate-low accept (13)	Moderate-high reject (26) Moderate-high discuss (26) Moderate accept (21)
Experiment	Low reject (6) High discuss (33) High accept (31)	Moderate reject (18) Very high discuss (59) Moderate-high accept (23)	Low reject (2) Very high discuss (47) Moderate-low accept (14)

Note. Scale to determine response levels was low (6 and under), moderate-low (7-14), moderate (15-22), moderate-high (23-30), high (31-38), very high (over 39).

In the control group, teams generally had higher levels of *reject* responses with lower levels of *discuss* and *accept* responses. On average, the highest level of responses was reject, followed by discuss, with accept responses at the lowest level. Teams in the control had moderate to high levels of reject responses; there are several reasons that explain why reject responses were the most common communication response in this treatment. First, teams struggled to keep their powders straight, which led to some reject responses:

S1: Now which one are you doing? This is baking soda.
S2: No, it's not.
S1: Yes, it is.
(*Team C2 Conversation, 364:366*)

S1: Can you stop making fun of my baking soda?
S2: It's not baking soda. It's sugar.
S1: Oh, it is? Wow.
(*Team C3 Conversation, 753:756*)

Second, teams disagreed about the names of the materials and how the powders were reacting to the tests:

S1: Oh, it's an oil lamp.
S2: No, it's an alcohol lamp.
S1: Whatever.
(*Team C1 Conversation, 281:283*)

S1: It's shrinking.
S2: It can't shrink unless it's melting.
S3: It's shrinking.
S2: It can't shrink.
S3: It was up to there and now it's down to there.
S2: You can't shrink it unless it's melting.
(*Team C3 Conversation, 627:632*)

Third, teams disagreed over how to execute certain directions; teammates would try to direct each other to do the activity a different way or simply get aggravated by how the teammate was performing the action.

S1: Hang it over the edge then.
S2: No.
S1: Yeah.
S2: No, I'm just going to hold it like this. You'll ruin it.
S1: No, that's the way to do it. You hang it over the edge.
S2: I'm just going to hold it like this the whole time.
(*Team C2 Conversation, 428:433*)

S1: Okay, but should I keep doing it?
S2: Stop, stop!
S1: Okay.
(*Team C3 Conversation, 238:240*)

S1: Whoa, whoa, whoa.
S2: Whoa there, you're going to spill it.
S3: Nobody fight.
(*Team C1 Conversation, 305:307*)

Finally, there were disagreements over dividing up responsibilities:

S1: Okay, I'm doing sugar.
S2: No, I already told you.
S3: No, he already said it. You can't ….
S1: No, you said burning.
S3: No, he's doing sugar.
(*Team C2 Conversation, 582:586*)

S1: No. Here, I'll give you this. I'll do this. Give it. Give me it. Give me; give me the thing.
S2: No. Let me do it for a little bit.
(*Team C3 Conversation, 1197:1198*)

When reject responses accumulated, team members would resort to criticism, sarcasm, and the use of insults such as "shut up," "you are terrible," "chicken," and "stupid." This, in turn, made some team members defensive:

S1: Now they're all mixed together. Good job (*observer comment: sarcasm*).
S2: No, they're not. I did it separately.
(*Team C2 Conversation, 1152:1563*)

S1: Job one, what is that?
S2: That's powdered sugar. I'm not stupid. I know what I'm talking about.
(*Team C1 Conversation, 307:308*)

Overall, working as group seems to be a source of frustration for control teams. Some teams even actively complained about having to work together:

Oh, my God, could someone like kick him out of the group please?
(*Team C2 Conversation, 1410*)

Okay, doing a group is like the most annoying thing ever.
(*Team C2 Conversation, 1402*)

Teams in the control had moderate levels of *discuss* responses. Since it was a hands-on experiment conducted in a small group environment, this level of discuss responses was expected. It was no surprise that a lot of the discuss responses revolved around observations of the experiments themselves:

S1: It slides and then it just goes right through.
S2: Whoa, wait. It's like hard now. Look at it. It turned hard.
(*Team C1 Conversation, 220:221*)

S1: Is it still a powder?
S2: Yes.
S1: Like how the other one turned into a solid?
S2: It kind of turned into a solid. This one kind of turned into vapor.
(*Team C2 Conversation, 896:899*)

S1: What do you see?
S2: Absolutely nothing.
S1: A piece of paper with some white stuff on it?
S3: All right, okay.
S2: Wait, wait, wait.
S1: Soggy piece of paper.
S2: That one's turned green.
(*Team C3 Conversation, 496:502*)

In the control, teammates also clarified with each other about what they were supposed to do:

S1: Wait, how many drops do you put on?
S2: I don't know.
S3: Three.
(*Team C2 Conversation, 258:260*)

S1: Where do we write that?
S2: Middle of the box.
(*Team C3 Conversation, 10:11*)

S1: Okay, so how are we going to describe that?
S2: Put powder into liquid and see what happens.
S1: No, it's too simple.
S3: There's more than one powder.
(*Team C1 Conversation, 17:20*)

Overall, teams in the control had moderate to low levels of *accept* responses which was notably lower than their moderate to high levels of *reject* responses. In contrast, teams in the experiment generally had higher levels of *discuss* and *accept* responses with lower levels of *reject* responses. Specifically, in each team, the highest level of responses was discuss, followed by accept, with reject at the lowest level. Teams in the experiment had high or very high levels of discuss responses; there are several reasons that explain why discuss responses were the most common communication response in this treatment. First, since the students all had different roles with unique information, they had to ask each other a lot of questions for clarification:

S1: What does yours say?
S2: Hmm...Doesn't say anything. It's empty.
(*Team E3 Conversation, 301:302*)

S1: Did we do the baking soda yet?
S2: Yeah... No, I did the sugar.
S1: I did sugar and cornstarch. Did you guys have baking soda yet or no?
(*Team E2 Conversation, 544:546*)

S1: Okay. And do you have any quests? Okay.
S2: What did you learn from [*game character's name*]?
(*Team E1 Conversation, 117:118*)

Second, in order to come to agreement about their understanding of game content, teams would discuss information collectively by challenging and clarifying pieces of information:

S1: Should we put the secretary or not?
S2: She was baking cookies …. ….
S1: Yeah.
S3: The secretary because she needed flour?
S2: Because she was cooking.
S3: In the kitchen?
S2: And money was stolen, and there was flour left.
(*Team E3 Conversation, 483:487*)

S1: So you guys have the janitor and [*game character's name*]?
S2: And money for the trip?
S1: Yeah, money for the trip.
S2: What trip?
S1: The trip for, to go to Pittsburgh.
(*Team E2 Conversation, 361:367*)

S1: The only one that fizzed was this, the baking soda. The other one didn't fizz.
S2: So the vinegar?
S1: The cornstarch didn't fizz.

S3: Cornstarch did not.
S1: Did not fizz.
S4: What about the iodine?
S3: It turned orange and looked like fleshy ….
(*Team E1 Conversation, 467:473*)

Third, teams sometimes discussed general ideas surrounding the game incident:

S1: Do you think the principal is actually going to be the thief?
S2: No, of course not. Okay, you actually think it's going to be the principal?
(*Team E2 Conversation, 30:31*)

S1: What is the white powder supposed to be?
S2: I think I know what it is.
S3: What if it's just like that sugar stuff, that white sugar?
(*Team E2 Conversation, 120:122*)

S1: I think the secretary …. the powder, but not the money.
S2: But then Destiny is using the powder and the money.
(*Team E1 Conversation, 183:184*)

Teams in the experiment had moderate to high levels of *accept* responses on average. Specifically, the below average team had a moderate-low level, the average team had a moderate-high level, and the above average team had a high level of accept responses. Since accept responses were stratified by achievement level, it seems that accepting responses may be based on standardized test achievement, or possibly, capacity to understand the science content. Here are some examples of interactions that were coded as *accept*:

1. The second speaker agrees with and supports the proposal of the original speaker.

S1: Could we say that the janitor's need for money for surgery, but has no connection to the powder?
S2: Yeah.
(*Team E1 Conversation, 208:209*)

2. The first speaker makes a proposal; the second speaker agrees and offers a next step.

S1: Okay, we have to go to the woodshop.
S2: All right, and the art room.
(*Team E1 Conversation, 422:423*)

3. The first speaker assists with documenting a proposal and the second speaker agrees.

S1: And the pH test is 8 by the way.
S2: Yeah.
(*Team E1 Conversation, 641:642*)

Overall, teams in the experiment had moderate to low levels of *reject* responses which was notably lower than their moderate to high levels of *accept* responses and their high or very high levels of *discuss* responses. The one team that had a moderate level of rejecting responses had a personality conflict with one team member; it was an ever-present problem that seemed to be minimized during the gameplay but was exacerbated during the classroom experiment portion. Even though rejecting responses were at a moderate level for this team, they had a moderate-high level of accepting and a very high level of discussing responses so the pattern still holds with rejecting responses at the lowest level of the three categories.

In conclusion, teams in the experiment and teams in the control showcased different patterns of communication responses. First, teams in the experiment had moderate to low levels of *reject* responses, while teams in the control had moderate to high levels of *reject* responses. Second, teams in the experiment had moderate to high levels of *accept* responses; while teams in the control had only moderate to low levels of *accept* responses. Lastly, teams in the experiment had high or very high levels of *discuss* responses; while teams in the control had only moderate levels of *discuss* responses.

In addition, Barron (2003) categorized accept and discuss responses as *engaged* responses, while reject responses are considered *non-engaged* responses. Teams in the experiment produced a fairly high level of engaged responses in comparison to their non-engaged responses. In contrast, teams in the control produced a fairly high level of non-engaged responses in comparison to their engaged responses.

RQ4: Scientific practices. The fourth research question is: How do scientific practices of teams in the INPLACE mobile game compare to those of teams in the control activity? In order to answer this question, a data analysis of the scientific practices among game teams and control teams will be presented. First, for each scientific practice, this section will summarize the patterns observed in the control teams followed by the patterns of game teams. Each practice will then be briefly summarized with a cross-case synthesis. Lastly, a complete synthesis of all practices will be presented as the final finding for the research question.

As described earlier, the scientific practices that occurred in team conversations were coded to align directly to the scientific practices from the NRC (2012a). When students discussed what was known about the investigation or tried to determine what needed to be answered, then the dialogue was coded as *Practice #1- Defining the Problem*. When students discussed their investigation plan or what information they needed to record, then the dialogue was coded as *Practice #3- Planning the Investigation*. When students discussed characteristics of the experiments they were observing, then the dialogue was coded as *Practice #4 – Interpreting Data*. When students tried to explain the relationships between data, then the dialogue was coded as *Practice #6 – Constructing Explanations*. When students supported or refuted an argument by citing relevant evidence, then the dialogue was coded as *Practice #7 – Arguing with Evidence*.

Practice #1: Defining the Problem. In the control group, each case study revealed a different level of Practice #1 according to the occurrences coded in the transcripts. For a summary of all the occurrences of this practice within each team, please refer to Table 23. Team C3 had only one occurrence coded as Practice #1; however, the other teams in the control group had some dialogue about this practice.

Table 23

Content-Analytic Summary Table of All Cases for Practice #1 - Defining the Problem

	Above Average	Average	Below Average
Control	Moderate (6)	Low (3)	Low (1)
Experiment	Moderate (5)	Moderate (6)	High (11)

Note. Scale to determine response level was low (1-4), moderate (5-8), high (9-14), very high (over 14).

How this practice was revealed in conversation took two forms. First, when students discussed the constraints or specifications needed for a solution, it was coded as Practice #1. This was the more common variety of this practice. However, there was another important—yet less pervasive—aspect to this practice as well. Asking questions is another component of Practice #1. Student dialogue was also coded as Practice #1 when students discussed their hypothesis, or more specifically, empirically answerable questions. Each aspect of Practice #1 will be treated separately in this summary.

Conversational dialogue in teams C1 and C2 defined the problem. In C1, after a few exchanges, the team finally summarized their thinking with this statement, "We will test three powders, do various experiments to ultimately conclude which powder is which" (*Team C1 Conversation, 41*). Team C2 came to a similar conclusion except they included more details:

> S1: In this space, a description of the experiment, and the experiment we will be doing is, the experiment we will be doing, trying to identify this new powder by experimenting on other known powders. It's in three tests, heat, what's the other ones?
> S2: pH, vinegar and.
> S3: Heat test.
> S2: Oh, that's it. I didn't know what the first says.
> S3: Heat test, pH and then.
> S2: No, there's iodine, too.
> S4: What's iodine? Oh, yeah, I remember.
> S2: The blue stuff, right?
> S4: Okay, let's write that down.
> (*Team C2 Conversation, 6:14*)

Team dialogue in C1 and C2 also revealed brief discussions of their hypotheses. In team C1, there was one brief exchange, which revealed little understanding of how to ask an empirically answerable question:

> S1: Smells like marshmallows again. Well how are we supposed to do this if we don't even see the powder? Like we can't guess what it is.
> S2: It's a hypothesis.
> S3: You have to make a hypothesis.
> S1: So what?
> S3: Just guess.
> (*Team C1 Conversation, 742:746*)

In team C2, there was also one brief exchange, which reveals a slightly better understanding of how to ask an empirically answerable question:

> S1: …. this one look? Apply vinegar to the mystery powder. Then it will either turn into dough or …. If it is sugar, it may …. the vinegar. If iodine, iodine, it will either turn black and brittle or turn yellow into dough-like substance. The mystery powder …. [mumbles]
> S2: I'm just going to say it will be between 9 and 2.
> S1: Okay. I said if I heat the mystery powder, then the mystery powder will smell like burned marshmallows and maybe turn black into a solid.
> (*Team C2 Conversation, 1455:1457*)

Overall, in the control group, defining the problem (Practice #1) had a low level presence with a low level of understanding revealed. Similar to the control, the experiment teams revealed varying levels of this practice. Interestingly, while the below average control team (C3) only utilized Practice #1 once, the below average experimental team (E3) had the highest level of Practice #1 of all cases in the study. To begin, team E3 had an ongoing conversation about the incident that occurred at the beginning of the game and worked together to develop their communal understanding of constraints of the problem. Here is a sample occurrence of this aspect of Practice #1:

> S1: …. cash register seems empty. The cash register is empty. …. …. mystery powder. …. …. All right, so now we have four suspects. [cross-talk] …. suspects.
> S2: There was a theft in the cafeteria.

S1: Yeah.
S2: Money was stolen.
S1: Oh, yeah. Some money was stolen. Also, what happened to the spray bottle?
S3: Money was stolen and there was damage done.
(*Team E3 Conversation, 440:446*)

Team E1 in the experiment also had similar style conversation as team E3. Here is a sample conversation where E1 works together to develop their description of the incident:

S1: Description of incident. I'm …. a space. Money stolen from cash register.
S2: From where? Where is this cash register?
S1: Cafeteria. Cafeteria. Mystery powder.
S2: Oh, I said power instead of powder.
S1: Mystery powder left near register. No forced entry. Therefore that's going to lead us to think that it is someone in our school. Okay, we each have the suspect. No forced entry. I don't think the penny was found next to the cash register. That's just to show you how much powder there is. People do that a lot.
S3: Actually really?
S2: Yeah, that's what it says.
S1: The cash register.
(*Team E1 Conversation, 80:88*)

The discussion among team E2 was not as detailed as the others but still revealed the essential information about the incident. Comments found in E2 included "There was a theft at the cafeteria"(*Team E2 Conversation, 45*) and "Some sort of suspicious powder was involved" (*Team E2 Conversation, 48*).

Team dialogue in E1 and E2 also revealed brief discussions of their hypotheses. While there was no dialogue about hypotheses among the below average control team (C3), the below average game team (E3) had an extended conversation about their hypotheses revealing a deeper understanding of empirically answerable questions. Here is a sample from the conversation that took place in team E3:

S1: All right, so if you add vinegar to the mystery powder, then blank, what will happen? All right, so what would happen if we add? Okay, this is like basically saying what's your hypothesis on if we add vinegar to the baking soda, to the mystery powder. So what would happen if we add vinegar to the mystery powder, judging on what we found

> out about these three things that were left at the crime scenes? So do you think it would sit on top, dissolve, react, or react or like bubble?
> *(Team E3 Conversation, 876)*

The other teams both had brief conversations about their hypotheses. Neither team revealed as deep an understanding as team E3 had showcased. In E1, the girls understood that you determine your hypothesis before you do the test.

> S1: Okay, now you write down the evidence summary and the hypothesis questions.
> S2: Are you going ….?
> S1: One second.
> S2: We don't do that until we test ….
> S1: Oh, okay…Hah! You have to fill it all out and then do it.
> (*Team E1 Conversation, 540:545*)

In E2, while trying to review his hypotheses with his team, one boy made an attempt to discuss his thoughts: "It will be black. It will melt and it will be…" *(Team E2 Conversation, 1209)*.

Overall, in the experimental group, defining the problem (Practice #1) had a moderate presence compared to the low presence of this practice in the control. Also, in reviewing student dialogue, the game teams revealed a stronger understanding of describing the problem as well as some understanding of how to create a hypothesis.

Practice #3: Planning the investigation. In the control group, teams showcased a high level of *Practice #3* according to the occurrences coded in the transcripts. For a summary of all the occurrences of this practice within each team, please refer to Table 24. When students discussed their investigation plan or what information they needed to record, then the dialogue was coded as *Practice #3- Planning the Investigation*.

Table 24

Content-Analytic Summary Table of All Cases for Practice #3- Planning the Investigation

	Above Average	Average	Below Average
Control	High (13)	High (13)	High (13)
Experiment	Low (4)	Low (4)	Low (2)

Note. Scale to determine response level was low (1-4), moderate (5-8), high (9-14), very high (over 14).

Team C1 had a discussion about their plan that initially revealed a very holistic understanding: "...we have to write an experiment. Like what are we going to do with …" (*Team C1 Conversation, 93*). They eventually boiled down their ideas into executable steps: "Oh, like for number one, when it says gather observations and stuff, we can like say we can look at, like we can observe the powders" (*Team C1 Conversation, 120*).

Team C2 achieved a basic, but solid understanding of their investigation plan: "Okay, using the first powder we can get, we will do the four tests - heat, pH, vinegar and iodine - and then keeping them in that order" (*Team C2 Conversation, 32*). Much later in the activity, they even referenced back to the plan: "Now we pick a mystery powder and test it like with all the tests" (*Team C2 Conversation, 998*).

Team C3 had some discussion about their investigation plan, but they had a lot less confidence in their ideas then teams of higher achievement levels:

> S1: ...pH test strips, and then go to whatever cornstarch, or sugar or baking soda, and then go to either an iodine or the vinegar test. See which one it combines.
> S2: So with vinegar first, all right. We're doing the vinegar test one first. …. that they all have reactions, something ….
> S3: I'm not sure.
> S2: Let's go with the vinegar test first, and then we'll go with the pH.
> S3: Okay.
> (*Team C3 Conversation, 120:124*)

Overall, in the control group, planning the investigation (Practice #3) had a high level presence with an appropriate level of understanding revealed. In contrast to the control, the experiment teams revealed low levels of this practice. In team E3, the dialogue never went beyond the game narrative; there was no real discussion of what steps they would need to execute in order to identify the mystery powder. Here is a sample from their conversation:

> S1: We have to finish this first though, the paper. All right, it says write in your group step by step will perform the investigation for the powder mystery. So the first step, so put number one, and then is asking the recent suspects.

S2: The first thing is to interview the suspects.
S1: Yep. The second part is, wait. All right, here. The first part is going to the scene of the crime, going to the crime scene. Going to crime scene is the first one. The second one is interviewing the recent suspects. All right. Yeah.
S3: The first one is to go to the crime scene?
S1: Yes. Do you have that for the first one? Okay, good. And the second one is interviewing the recent suspects. Then come, wait. Go through the evidence and find; go through the evidence and find a prime suspect other than the suspect, other than the people we interviewed.
S2: So go through the evidence.
S1: Okay, go through the evidence and then find a prime suspect.
(*Team E3 Conversation, 519:526*)

Team E1 also planned out their investigation in terms of where to go in the game:

> S1: Okay, plan out in this space, write the steps that you'll perform to identify the mystery powder. Well let's just say go talk to the janitor. (*Team E1 Conversation, 89*)

However, unlike E3, team E1 eventually came to the understanding that they needed to conduct tests on powders:

S1: All right. You guys are only doing pH tests? I'm doing a vinegar test. [cross-talk] Olivia, what test are you doing? [Inaudible] What?
S2: Heat test.
S1: Heat test. And, …., what are you doing?
S3: Iodine test.
S1: Iodine test, okay.
(*Team E1 Conversation, 221:225*)

Team E2 had the best understanding of what they would need to do to identify the mystery powder. They were able to bring together the real science content with the game narrative. This team achieved this level of understanding with a little bit of teacher scaffolding:

> Like if I take the powder from the scene and see if it burns or melts, say it melts, and then like so all our suspects have somehow been involved with a powder, but they're not all the same. So whichever one. Or it could also; it couldn't just be one. It could be two powders. That narrows it down a little bit and then you figure it out from there. So if you find the person that their powder melts and then the one at the scene melts, then that's like …. (*Team E2 Conversation, 615*)

Overall, teams in the experiment had a low presence of planning the investigation (Practice #3). The discussions they had were tainted by the game narrative; while not a surprising finding, it was an unfortunate finding. Observations of teams planning their investigation revealed that this written assignment came at a very bad time during gameplay. Teams had just unlocked Chapter 2, their first major accomplishment in the game, and they were almost immediately asked to write down their investigation plan. When students were engaged in the game and then asked to write in the packet, they vocalized complaints such as "I hate the darn packet" (*Team E1 Conversation, 532*).

In conclusion, in the experimental group, Practice #3 had a low presence compared to the moderate presence of this practice in the control. Additionally, in reviewing student dialogue, teams in the control had a better understanding of the plan they needed to execute in order to determine the identity of the mystery powder. Although, one team in the experiment did show potential for marrying the game narrative with the science content, so a game re-design could improve game players understanding of planning the investigation (Practice #3).

Practice #4: Interpreting data. In both the control group and the experimental group, all case studies revealed a very high level of Practice #4 according to the occurrences coded in the transcripts. For a summary of all the occurrences of this practice within each team, please refer to Table 25. There are two main reasons why Practice #4 was the most common scientific practice in both groups. First, teams spent the large majority of their time conducting hands-on experiments and making observations. Second, the bulk of the written work included filling in two data tables. Therefore, an enormous amount of dialogue pertained to data characteristics and observations. While both control teams and game teams had a lot of dialogue coded to Practice #4, there were differences in how they talked about their observations and data.

Table 25

Content-Analytic Summary Table of All Cases for Practice #4 – Interpreting Data

	Above Average	Average	Below Average
Control	Very high (59)	Very high (106)	Very high (65)
Experiment	Very high (21)	Very high (38)	Very high (30)

Note. Scale to determine response level was low (1-4), moderate (5-8), high (9-14), very high (over 14).

In the control group, teams had a lot of short, brief observations. Many of their comments were very basic, lacking precise details. Their choice of words was sometimes comical. Here are some samples from all levels:

> It eats, like eats away the sugar. (*Team C1 Conversation, 212*)
> It's like turning funnel-shaped at the bottom. (*Team C1 Conversation, 469*)
> Like yellow snow. This turned into like dough. (*Team C2 Conversation, 270*)
> It looks like it's alive. (*Team C2 Conversation, 291*)
> It looks like boogers and snot. (*Team C2 Conversation, 405*)
> It's rising. (*Team C2 Conversation, 830*)
> It's sliding down. (*Team C2 Conversation, 1199*)
> It just sizzled …. (*Team C3 Conversation, 243*)
> Ooh, it got all hard. (*Team C3 Conversation, 337*)
> It's turning black. (*Team C3 Conversation, 352*)
> Nothing is happening. (*Team C3 Conversation, 481*)
> That smells great now. That smells great. (*Team C3 Conversation, 612*)

Additionally, teams in the control group were not always specific about the particular powder with which they were working. Here is a selection of examples from all levels:

> S1: Ooh, this one breaks.
> S2: This one got hard.
> S3: Ooh, it's like jelly.
> (*Team C1 Conversation, 394:396*)
>
> This one turned yellow again, and then this one turned like different colors.
> (*Team C1 Conversation, 660*)
>
> It is. That's black and that one's brown.
> (*Team C3 Conversation, 375*)
>
> Whatever one was right there. It's turning green.
> (*Team C3 Conversation, 507*)

S1: Ooh, what the heck? What are you doing? …., which one is that? Oh, yeah, I'm doing cornstarch.
S2: Why?
S1: Because that's what I said before and it turns black.
(*Team C2 Conversation, 522:524*)

S1: Which one is this right now? The one that we're doing.
S2: That's the vinegar.
S1: No, I mean this. It's cornstarch, right?
S2: It's cornstarch.
(*Team C2 Conversation, 107:110*)

Since the teams were generally basic in their descriptions and were not always specific about the particular powder that they were testing, it led to some confusion among team members:

S1: Which one's the one that looked like brown, or black or gray?
S2: Cornstarch. Baking soda and cornstarch. Wait, which one is …. with baking soda and cornstarch ….
(*Team C3 Conversation, 1084:1085*)

S1: You remember which one's which?
S2: Yes, this one's baking soda. This one's; what's the one that you just did?
S3: This one? This one is the red one. That's sugar.
(*Team C2 Conversation, 459:461*)

S1: What is this?
S2: Sugar.
S3: No, it's baking starch.
S2: It's soda.
S3: It's baking soda. I'm positive that it's baking soda.
S1: B.
S3: Like I said, baking soda. I'm not stupid. I told you that it's baking soda.
(*Team C1 Conversation, 191:197*)

Overall, in the control group, there was a very high level of presence of interpreting data (Practice #4); however, team comments lacked detail which sometimes led to confusion. In the experimental group, there were was also a very high level of occurrences of this practice, yet numerically speaking, the game teams had fewer occurrences than the control. Although one might say this means the practice had less presence, the team comments were fewer in number but higher in substance. To begin, observational comments in game teams included descriptive wording and specified the particular powders:

> The cornstarch burned and turned brown and colored, and then the sugar melted and it looked like yellow acid. (*Team E2 Conversation, 668*)
>
> This is the baking soda one. So for that one it reacted and started to bubble. (*Team E3 Conversation, 802*)
>
> S1: Gets all like liquidy. That's why I'm not sure that ….
> S2: So wait. It like, it turns yellow?
> S3: This is the iodine test for sugar.
> S2: It turns yellow and then it turns liquidy and white again.
> (*Team E2 Conversation, 720:723*)

Rather than observations being shouted out individually as they often were in the control teams, game teams used group discussion to obtain a nuanced observation:

> S1: It is definitely burning.
> S2: Or is it caramelizing?
> S1: No, it's burning.
> (*Team E1 Conversation, 616:618*)
>
> S1: Look at this one. Look at this one. It changes the color. I know that for sure. But it doesn't really change. Do you think it changes ….?
> S2: It does. That looks like a paste.
> S1: Let's watch it again. Excuse me. I don't think it's the color.
> S2: Yeah, it does. See it's more like a paste now. See?
> S3: Wait.
> S1: Yeah, but it's still …. …..
> S2: It changes ….
> S3: Do it again. Do it one more time, please. Please.
> S1: I know for sure it changes the color. But here it's still got that little like rattling sound….
> S3: Yeah.
> S1: …that you would have, that's how you'd like …. …. Plus it doesn't change the texture.
> (*Team E2 Conversation, 893:901*)
>
> S1: All right, for that it dissolved and turned it yellow.
> S2: Yellow?
> S1: Yeah, look at it. Actually it turned it orange.
> S2: Yeah, …. turned orange.
> S1: All right.
> (*Team E3 Conversation, 806:810*)

S1: And that changed color, right? That changed color, right?
S2: No, it just like sat on top. It sat on top. I never really observed anything.
S1: Yeah, but it changed the color of it, of some of it.
(*Team E3 Conversation, 729:731*)

In some game teams, observations naturally led to comparisons.

S1: Yes. Baking soda fizzed. Well, wait, I don't think the other one; I don't know if the other one fizzed or not. Here, let me try.
S2: The sugar and the cornstarch …. The baking soda ….
S3: What about the iodine?
S2: Or the vinegar.
S1: One second.
S2: Did it fizz?
...
S1: Okay, I'm sorry. I lied about the other thing before I think. The only one that fizzed was this, the baking soda. The other one didn't fizz.
(*Team E1 Conversation, 460:467*)

S1: See. It changes the color and the texture of the sugar.
S2: That changes the color. It's black.
S1: Yeah, but when …. it doesn't do anything to the …. [cross-talk] Like literally it just stays the same, like as opposed to sugar, …. the sugar one.
(*Team E2 Conversation, 655:657*)

Okay. Well the first one, for the cornstarch it does fizz, and then for the sugar it does not fizz. (*Team E2 Conversation, 663*)

Finally, the comments of certain team members seemed to reveal expectations of how the powder would perform, or that their expectations had not been met:

The mystery powder, to me it looks kind of like, it looks like cornstarch and it's probably just going to stay on top. (*Team E3 Conversation, 876*)

S1: Why is it hardening? Strange.
(*Team E2 Conversation, 1175*)

S1: It's funny how it was rising.
S2: I thought it was going to shoot out ….
(*Team E2 Conversation, 1325:1326*)

In conclusion, dialogue surrounding interpreting data (Practice #4) was more substantive and specific in the experimental group as compared to the team dialogue in the control group.

Additionally, in reviewing student dialogue, teams in the control had more confusion due to their lack of substance and specificity.

Practice #6: Constructing explanations. In the control group, C1 and C2 teams had a moderate level of Practice #6, while team C3 had a low level. For a summary of all the occurrences of this practice within each team, please refer to Table 26.

Table 26

Content-Analytic Summary Table of All Cases for Practice #6 – Constructing Explanations

	Above Average	Average	Below Average
Control	Moderate (6)	Moderate (7)	Low (3)
Experiment	High (9)	High (14)	Low (4)

Note. Scale to determine response level was low (1-4), moderate (5-8), high (9-14), very high (over 14).

The below average control team (C3) had very little evidence of constructing explanations. This sample discussion exemplifies the best example of their capacity to explain the relationships between the data.

> S1: What did you write?
> S2: I wrote so our conclusion is that it is cornstarch and baking soda because all the tests had something to do with the cornstarch and baking soda, like the iodine, the vinegar and the pH test.
> (*Team C3 Conversation, 1241:1242*)

Team C1and C2 had a moderate level of this practice. The occurrences in these teams included some evidence of explaining relationships between data as well as creating explanations based on valid evidence obtained from their powder testing:

> S1: Three, it bubbled, I knew it. It didn't change color though.
> S2: Just like the baking soda.
> (*Team C2 Conversation, 1141:1142*)

> S1: "I think it's going to be sugar and baking soda, because it smelled, like you could smell it, and sugar was the only one you could smell, and it was just bubbling like the baking soda." (*Team C2 Conversation, 1149*).

S1: It raised?
S2: Yeah.
S1: See, oh cornstarch raised.
S2: And it smelled like marshmallows.
(*Team C1 Conversation, 955:958*)

Overall, in the control group, constructing explanations (Practice #6) had a moderate level presence with the average and above average teams; in their dialogue they exhibited important aspects of the practice. The below average team achieved only a low level of Practice #6; in their dialogue they exhibited the bare minimum of understanding of the practice.

In the experiment, the below average team (E3) also had a low level presence of this practice. Team E3 constructed some explanations for why certain characters could be considered suspects. They used evidence obtained from the game. For example, as the team discussed the secretary, a team member said, “She made the cookies…in the kitchen by using flour” (*Team E3 Conversation, 490*). This team did not construct explanations of any relationships between the data as it pertained to the powder evidence.

Team E2 also constructed explanations for why certain characters in the game could be considered suspects; however, they were also able to go beyond the game narrative. Team members were able to explain relationships between data:

S1: Did the baking soda fizz a lot more than the cornstarch?
S2: Here, let me check. So baking soda. Yeah, baking soda fizzed a lot more than; it fizzed a lot more than …. cornstarch.
(*Team E2 Conversation, 1343:1344*)

There were even able to take the scientific powder evidence and relate it back to the game suspects: "But does everything else match up with the janitor?" (*Team E2 Conversation, 1364*).

Just as the other teams constructed explanations about suspects, team E1 did as well. They were also able to discuss relationships between the data in ways similar to team E2:

S1: Oh, no, no, no, because this is fizzing and the only one that fizzed is baking soda.
S2: Okay, well, this is fizzing. The only one that fizzed is baking soda, so.
(*Team E1 Conversation, 581:582*)

The most interesting occurrence of this practice for team E1 happened when they were not even working with their own data. As the girls were working on their packet and classmates were starting to test the mystery powder, the girls started to comment on what other teams were doing and starting construct explanations about their observations:

S1: Oh, my God, that thing smells really good.
S2: Probably sugar. Probably sugar burning.
S1: Oh, yeah, you're caramelizing sugar.
(*Team E1 Conversation, 537:539*)

Overall, in the experimental group, constructing explanation (Practice #6) had a high level presence with the average and above average teams; their use of the practice traversed both game-related explanations and purely scientific explanations as they worked with the powder evidence, and even extended beyond both situations to explain a sweet smell that had taken over their classroom. The below average team achieved only a low level of Practice #6 because they only revealed game-related explanations.

To summarize, the below average teams in both the control group and the experimental group only achieved a low level of Practice #6. However, in closely examining the average and above average teams from both groups, a contrast emerged. In the experiment, teams achieved a *high* level of Practice #6, while comparable teams in the control only achieved a *moderate* level of Practice #6.

Practice #7: Arguing with evidence. In the control group, each case study revealed a different level of Practice #7 according to occurrences coded in the transcripts. For a summary of all occurrences of this practice within each team, please refer to Table 27. While C3 had two comments coded as Practice #7, they were prompted by teacher interaction; this team never exhibited this practice on their own. However, other teams had dialogue about this practice.

Table 27

Content-Analytic Summary Table of All Cases for Practice #7 – Arguing with Evidence

	Above Average	Average	Below Average
Control	Moderate (7)	Low (3)	Low (2)
Experiment	Moderate (6)	Moderate (5)	Moderate (5)

Note. Scale to determine response level was low (1-4), moderate (5-8), high (9-14), very high (over 14).

Team dialogue in C1 and C2 revealed some arguments based on evidence. In team C2, there were a few arguments made by one particular student. For example:

> That's what I said, because theirs smelled. Like you could smell the marshmallows in it when it was burning [cross-talk], and, what else, because then it bubbled when the vinegar touched it. (*Team C2 Conversation, 1238*)

In team C1, a few students had an argument-based conversation at the end of the activity. Here is a selection from that conversation:

> S1: Because it turned liquidy like the baking soda and it pumped up like sugar.
> S2: Yes, it did.
> S1: It didn't turn liquid. It just burned black. Or it burned.
> S2: It hardened in the iodine test.
> S3: It did?
> S2: Yeah, it turned black and hardened.
> S1: So maybe cornstarch and sugar.
> (*Team C1 Conversation, 939:945*)

Overall, in the control group, the presence of arguing with evidence (Practice #7) correlated with achievement level. First, the below average team (C3) did not showcase this practice at all on their own. Second, the average team (C2) revealed a few arguments; however, they were all made by the same student. Finally, the above average team (C1) had a conversation where different students made statements supported with evidence. Additionally, it was notable that this practice only occurred briefly at the very end of the activity when they were working on their final conclusion.

In contrast to the control, the game teams all revealed moderate levels of arguing with evidence (Practice #7). In the below average game team (E3), their exhibition of this practice revolved entirely around the game narrative, rather than the powder evidence. They used the practice early in the game when they were trying to determine who the suspects were. As the team discussed the secretary, one team member said: "So the secretary because she was; secretary because she was baking, she was cooking, was cooking in the kitchen" (*Team E3 Conversation, 489*). The team also exhibited this practice again at the end of the game as there were discussing the real thief: "So the janitor because he was at the school late and he was the only one here at the time of the theft, the theft happened" (*Team E3 Conversation, 1033*). Unfortunately, this team never used evidence-based arguments when talking about the powders.

Team E2 also argued with evidence about the suspects in the game: "He had the motive for money and he talked about like the baking was it, the white residue" (*Team E2 Conversation, 346*). Yet, unlike E3, team E2 eventually argued with evidence as they tried to explain what powders were present in the mystery powder:

> So...the cornstarch burned. So the only thing that burns is the cornstarch, so it has to be the cornstarch. (*Team E2 Conversation, 1351*)

In team E1, the girls had an extended discussion as they argued with each other for who the suspects were based on evidence they had collected in the game:

S1: …. the money, but.
S2: I think the secretary …. the powder, but not the money.
S3: But then Destiny is using the powder and the money.
S1: You really don’t put baking soda in cookies usually.
S3: Yes, you do.
S4: She said she used sugar.
S2: Yeah.

S1: …. bottle of baking soda …. [cross-talk]
S3: It could be powdered sugar.
S1: The same bottle that the girl used for her project, the rocket and...
(*Team E1 Conversation, 182:191*)

The girls utilized this practice again as they tried to reason out what powders were present in the mystery powder:

S1: All right, that's cornstarch.
S2: No, it's not just cornstarch.
S3: Because it's also all of these.
S2: It's all of the characteristics of everything. It fizzed, it's clear-watery, and it burned under fire.
(*Team E1 Conversation, 637:640*)

Overall, game teams had moderate levels of arguing with evidence (Practice #7). This was distinctively different from the control group where some teams did not even showcase without teacher intervention. Teams in the experiment also showcased this practice on more than one day of the activity whereas teams in the control only showcased it once at the end.

Scientific practices summary. In conclusion, game teams and control teams showcased different patterns of scientific practice during their conversations. For a summary of all the occurrences of these types of practices within each team, please refer to Table 28.

Table 28

Content-Analytic Summary Table of Scientific Practices for All Cases

	Above Average	Average	Below Average
Control	Interpreting data: Very high (59) Planning investigation: High (13) Arguing with evidence: Moderate (7) Constructing explanations: Moderate (6) Defining the problem: Moderate (6)	Interpreting data: Very high (106) Planning investigation: High (13) Constructing explanations: Moderate (7) Defining the problem: Low (3) Arguing with evidence: Low (3)	Interpreting data: Very high (65) Planning investigation: High (13) Constructing explanations: Low (3) Defining the problem: Low (1) Arguing with evidence: Low (2)

Experiment	Interpreting data: Very high (21) Constructing explanations: High (9) Arguing with evidence: Moderate (6) Defining the problem: Moderate (5) Planning investigation: Low (4)	Interpreting data: Very high (38) Constructing explanations: High (14) Arguing with evidence: Moderate (5) Defining the problem: Moderate (6) Planning investigation: Low (4)	Interpreting data: Very high (30) Defining the problem: High (11) Arguing with evidence: Moderate (5) Constructing explanations: Low (4) Planning investigation: Low (2)

Note. Scale to determine response level was low (1-4), moderate (5-8), high (9-14), very high (over 14).

In the control group, teams generally had very high levels of *interpreting data,* high levels of *planning the investigation,* varying levels of *defining the problem,* moderate to low levels of *constructing explanations,* and low levels of *arguing with evidence*. In the experimental group, teams generally had very high levels of *interpreting data,* moderate to high levels of *constructing explanations,* moderate levels of *arguing with evidence,* varying levels of *defining the problem,* and low levels of *planning the investigation*.

In reviewing the occurrences of each practice, a deeper analysis revealed that simply reviewing level of occurrences did not tell the whole story. First, for defining the problem (Practice #1), game teams revealed a stronger understanding of describing the problem as well as some understanding of how to create a hypothesis. While control teams did showcase this practice, they only revealed a basic understanding of describing the problem and a very basic understanding how to create a hypothesis. Second, for planning the investigation (Practice #3), teams in the control had a better understanding of the plan they needed to execute in order to determine the identity of the mystery powder than teams in the experiment. Third, for interpreting data (Practice #4), although both treatments had a high level of occurrences, teams in the experiment offered observations that were more specific and substantive than teams in the control. Fourth, for constructing explanations (Practice #6), the below average teams from both treatments struggled somewhat with this practice exhibiting only a basic understanding.

However, in comparing the higher achieving students, game teams constructed explanations about both the game narrative and the scientific content leading to more opportunities to showcase this practice whereas control teams only explained the science content. Finally, while arguing with evidence (Practice #7), game teams revealed their capacity to argue with evidence more than once during the activity; multiple team members were also involved in making evidence-based arguments. Not all control teams showcased this practice on their own; for those that did, they only revealed it once at the end of the activity and only one control team had multiple members exhibiting the practice. Other than for planning the investigation (Practice #3), game teams showcased a deeper ability to integrate the scientific practice into their conversations than control teams.

RQ5: Additional differences. The fifth research question is: How else are treatment groups different when qualitatively analyzed at the team level? In order to answer this question, qualitative data was analyzed from the ground up; three themes emerged including the nature of their non-scientific comments, group dynamics, and their teacher involvement. For the purpose of presenting the findings in this section, the themes will be reconfigured as (1) group dynamics and language style and (2) teacher involvement and the learning process. The first part will present findings about group dynamics and students use of commands versus communal language. The second section part will present findings about teacher involvement and students' confusion and off-topic discussions.

Group dynamics and language style. As mentioned earlier, some codes emerged during a second round of coding. The researcher noticed differences in the general language style of the treatment groups. Students in the control were frequently telling each other what to do. They were using language such as, "don't reach across the table like that—here—give it" (*Team C1*

Conversation, 510), "put the whole entire thing in" (*Team C2 Conversation, 122*), and "go get the other one" (*Team C3 Conversation, 285*). To capture this type of directive language, a new code was created called *commands*. Additionally, the researcher noticed that students in the experiment were addressing the team collectively, rather than one specific team member. They were also referring to the team as an entity with words such as "we," "we're," and "let's." To capture this type of communal language, a new code was created called *communal language*. For a summary of all the occurrences of these types of language styles within each team, please refer to Table 29.

Table 29

Conceptually-Ordered Summary of Group Dynamics and Language Style for All Cases

	Group Dynamics	Commands (occurrences)	Communal Language (occurrences)	Communication Responses
Team E1	Group discussed and selected game roles. Group made decisions together.	Low (6)	High (39)	High discuss High accept Low reject
Team E2	Group selected game roles. Not too many group issues during gameplay. Some issues in classroom.	Moderate (23)	High (63)	Very high discuss Moderate-high accept Moderate reject
Team E3	Everyone chose game roles; however, boy leader emerged. Very good teacher and communicator. No resentment towards leader.	Moderate (32)	High (43)	Very high discuss Moderate-low accept Low reject
Team C1	No leader emerged. Democratic but not effective.	Moderate (28)	Moderate (22)	Moderate discuss Moderate accept Moderate reject
Team C3	Girl leader emerged. Decent communicator and delegator. Boys followed willingly while at times resisting.	High (49)	Low (7)	Moderate-high discuss Moderate accept Moderate-high reject
Team C2	One boy took leadership. Ineffective communicator and very controlling. Teammates showed resentment.	High (50)	Low (17)	Moderate discuss Moderate-low accept High reject

Note. Scale to determine commands and communal language levels was low (19 and under), moderate (20-38), high (over 38). For scale and numeric occurrences of Communication Responses, please refer to Table 22.

Earlier findings have shown that teams in the experiment had moderate to low levels of *reject* responses which was notably lower than their moderate to high levels of *accept* responses and their high or very high levels of *discuss* responses. In contrast, teams in the control had moderate to low levels of *accept* responses which was notably lower than their moderate to high levels of *reject* responses. Those findings have been recreated in the conceptually-ordered summary shown in in Table 28. Earlier findings have also shown that game teams showcased a deeper ability to integrate scientific practices into their conversations than control teams.

Investigating the group dynamics, language style, and communication responses of each case reveals the effectiveness of team communication. This display is conceptually ordered to begin with the team that had the most effective communication pattern to the team that had the least effective communication strategy. The three game teams are at the top, while the three control teams are at the bottom. In this section, each case will be summarized according to their group dynamics, language style, and communication responses and supported with student quotations. Then, to conclude, a cross-case comparison will synthesize the differences in group dynamics and language style based on treatment group.

Team E1. This team had a low level of commands and a high level of communal language. Additionally, they had high accept and discuss responses and low reject responses. From the very beginning of the activity, the girls on this team were using communal language: I know we're going to, but like we'll have to figure it out for ourselves (*Team #4 Conversation, 43*). They asked each other a lot of questions, such as:

> What does everybody have on their screens now? (*Team E1 Conversation, 152*)
> "All right. You guys are only doing pH tests? I'm doing a vinegar test. [cross-talk] Olivia, what test are you doing?" (*Team E1 Conversation, 221*)

Their constant use of the term "we" was consistent throughout the activity. When one girl saw a quest pop up on her screen, she said, "we have a new quest" (*Team E1 Conversation, 480*). Beyond communal language, the team members even showed polite manners to each other:

> S1: Baking soda was like 8.
> S2: Thank you.
> (*Team E1 Conversation, 577:578*)

This team showcased low levels of commands, and when they did deliver a command, the language was fairly soft: "Wait, let me see the sugar test" (*Team E1 Conversation, 596*). All in all, this team performed extremely well as a group in terms of their team discussions and decision-making. The term for this group dynamics structure will be: *community action*.

Team E2. This team had a moderate level of commands and a high level of communal language. Additionally, they had moderate to high accept and very high discuss responses with moderate reject responses. The moderate level of reject responses confirms that this team did not always work well together, but they had high communal, moderate to high accept, and very high discuss, so they were able to communicate effectively and get the job done. Similar to the community action team, Team E2 asked each a lot of questions: "what do we do?", "can we do both?" , and "where are we going?" (*Team E2 Conversation, 212; 267; 969).* They also referred to their team as a distinct entity with specific goals to achieve:

> So we need to go to the principal. (*Team E2 Conversation, 21*)
> We need to fill out the form now. (*Team E2 Conversation, 264*)
> Okay. Then we have to press start two in the decoder. (*Team E2 Conversation, 373*)

This team used commands sporadically throughout the activity, both during gamplay ("You don't scan that. You have to scan your color.") and during the hands-on lab activity ("Give me the petri dish.")(*Team #5 Conversation, 69; 1096*). Overall, this team had a lot of democratic discussions that included enough support for each other's ideas that they could overcome the moderate amount of rejecting responses. The term for this group dynamics structure will be: *effective democracy*.

Team E3. This team had a moderate level of commands and a high level of communal language. Additionally, they had moderate to low accept and very high discuss responses with low reject responses. One boy seemed to grasp the technology quickly and emerged as a leader in the team who could help his teammates use the technology: "You have to scan the barcode….You see this barcode?" (*Team E3 Conversation, 147:148*). The leader was even helpful when students were completing their worksheets: "No, you put it under vinegar test because that was the video. That's a different test. You put it under here. We're just doing this, this and that." (*Team E3 Conversation, 711*). During the hands-on lab activity, he was a caring nurturing teacher to others in his team: "Take a couple drops and drop them into that, and watch, it stains clothes. Just like do three or four drops. Keep going. All right, that's good." (*Team E3 Conversation, 953*). Given this boy's method of instructional communication, this team had a moderate level of commands, but the tone of his language was comforting, not demanding. So this team had a high level of communal language because they were functioning as discrete entity: "Okay, so, well, we have to do all of them, so why don't we go to the woodshop first because it's right here." (*Team E3 Conversation, 584*). They had a very high level of discussion because the one leader boy included and expected everyone to participate. The team had a low level of reject responses because the leader's ideas were rarely rejected. However, there were also rarely supported so there was a moderate to low level of accept responses. The term for this group dynamics structure will be: *expert guide*.

Team C1. This team had a moderate level of commands and a moderate level of communal language. Additionally, they had moderate accept and moderate discuss responses with moderate reject responses. They shared some similar language patterns to the game teams: (1) they referred to themselves as a discrete entity: “We have to test. We have to put the powders in this liquid and see how they react.” (*Team C1 Conversation, 16*); (2) they were a team with specific goals to achieve: “We have to make observations.” (*Team C1 Conversation, 97*); and (3) they asked questions to the group: “How do we get that done?” (*Team C1 Conversation, 62*). However, in contrast to the effective democracy, they only had moderate communal, moderate accept, and moderate discuss responses, so they did not have enough of these language styles to be effective. The moderate level of reject responses confirms that this team did not always work well together. The moderate presence of commands confirms that there were times that members made directive comments to individuals rather than collectively to the team:

> S1: That’s why I told you to hold it longer.
> S2: [cross-talk] …. It just turns black.
> S1: You have to keep doing it.
> (*Team C1 Conversation, 493:495*)

Overall, there was not enough communal language to overcome the commands and rejecting responses, so they were not as competent as the effective democracy. Therefore, the term for this group dynamics structure will be: *ineffective democracy*.

Team C3. This team had a high level of commands and a low level of communal language. The girl took the lead role and directed the boys; therefore, there was a high level of commands. Here are a few examples:

> Then go get some cornstarch, sugar and some baking soda. (*Team C3 Conversation, 58*)
> No, put more on. (*Team C3 Conversation, 340*)
> You do this one. (*Team C3 Conversation, 688*)
> Take it off the flame. Look at the side of it. (*Team C3 Conversation, 708*)

Unlike the expert guide who took a more communal approach to his instructions, this leader communicated with commands but did not include a lot of communal language and sometimes got aggravated with her team members. For example:

> S1: No, stop.
> S2: Don't point it. This is as terrible ….
> S1: Give it to me.
> S3: Give it to me.
> S2: Don't point it at anybody.
> (*Team C3 Conversation, 607:611*)

Additionally, they had moderate accept and moderate to high discuss responses with moderate to high reject responses. The leader included the others in the procedures, so there was moderate to high level of discuss responses. There was a relatively even amount of accepting and rejecting responses so team members supported and rejecting each other's ideas equally. In other words, while the group dynamics were not ineffective, they were not effective either. The low communal language and moderate accept responses showed that there was not enough engagement with each other's ideas. Since the team members were all taking part in the action, but did not communicate communally, the term for this group dynamics structure will be: *egalitarian leadership*.

Team C2. This team had a high level of commands and a low level of communal language, similar to the egalitarian leadership. The lack of communal spirit was evident from the beginning of the activity:

> S1: Okay, go. Someone go.
> S2: Someone else.
> S1: Anyone? Anyone at all.
> S3: I'm going.
> S2: A sacrifice.
> S1: Oh, my God.
> S2: Just kidding.
> S1: Are you kidding? Are you sure?
> S4: Go and get it ….
> (*Team C2 Conversation, 62:70*)

It was not long before one boy took the lead role and directed the others; therefore, there was a high level of commands such as “Hey, I told you to do it. Put it on the like top of the ….” (*Team C2 Conversation, 439*) and “Take it out. Take it out, Brittany. Take it out.” (*Team C2 Conversation, 568*). The communal language generally came from the one girl in the team who was trying to help keep everyone on track. Here are a few examples:

> Okay, we’re going to write that, right? (*Team C2 Conversation, 22*)
> We’re supposed to clean up. (*Team C2 Conversation, 133*)
> We have to take it out and put what number it is. Right? (*Team C2 Conversation, 469*)
> Okay, let’s just do the work now. (*Team C2 Conversation, 1426*)

The power struggle between the leader boy and this girl led to a high level of reject responses and even a number of insults, such as:

> Oh, my God. You are so stupid. (*Team C2 Conversation, 551*)
> Because he’s not that bright. (*Team C2 Conversation, 797*)
> But he’s annoying. (*Team C2 Conversation, 1400*)
> You guys are clinically slow workers. (*Team C2 Conversation, 1556*)

While this team managed to have moderate level of discuss responses, they rejected each other’s ideas frequently and rarely supported each other. This made them less effective than the egalitarian leadership team. The fact that one boy took the leadership role and had to be in control at all times makes the term for this group dynamics structure: *dictatorship*.

Overall, teams in the experiment not only had higher levels of accepting and discussing responses but they also had higher levels of communal language and lower levels of commands. These patterns of team communication seemed to connect with better group dynamics and more effective team communication. Teams in the control not only had higher levels of rejecting responses but they also had lower levels of communal language and higher levels of commands. These patterns of team communication seemed to connect with less effective group dynamics and poor team communication skills.

Teacher involvement and the learning process. As mentioned earlier, some codes emerged during a second round of coding. The researcher noticed that students in both treatment groups were having off-topic conversations and managing some level of confusion. Students in both treatments had conversations about their clothes, after-school activities, television shows, movies, and popular music—to name a few. To capture these types of off-topic conversations, a new code was created called *off-topic*. Additionally, the researcher noticed that students were verbalizing their confusion with questions such as "Are we supposed to do the next one?" and confirming their uncertainty with statements such as "I have no idea" (*Team C2 Conversation, 19:22*). To capture these comments pertaining to confusion, a new code was created called *confusion*. For a summary of all the occurrences of these types of comments within each team, please refer to Table 30.

Table 30

Conceptually-Ordered Summary of Teacher Involvement and Learning Process for All Cases

	Teacher Involvement	Confusion (occurrences)	Off-Topic Comments (occurrences)	Group Dynamics Structure
Team E3	Teacher assistance with technology. Scaffolding. Basic instruction. Brief reviewing. [NOTE: This group had a dedicated aide on the first day.]	Low (9)	Low (1)	Expert Guide
Team E1	Very low teacher involvement. No direct instruction, scaffolding, or review. Only assistance as needed.	Moderate-Low (16)	Moderate-Low (12)	Community Action
Team E2	In the halls, teacher assistance with technology and corralling students at end of periods. In class, some instruction and scaffolding. No reviewing.	Moderate (27)	Moderate (29)	Effective Democracy

Team C3	Direct instruction; scaffolding; reviewing	Moderate (27)	High (46)	Egalitarian Leadership
Team C2	Direct instruction; scaffolding; reviewing	Moderate-High (31)	Moderate-High (30)	Dictatorship
Team C1	Some instruction; some reviewing; assistance as needed	Moderate-High (36)	Moderate (24)	Ineffective Democracy

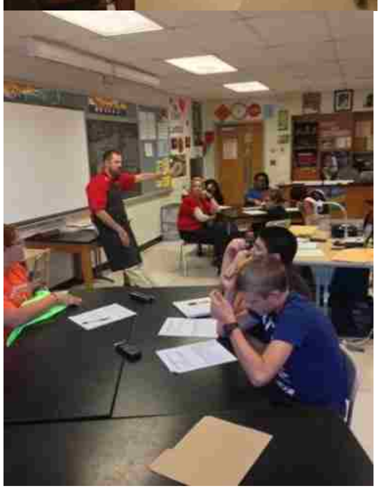
Note. Scale to determine levels of confusion and off-topic comments was low (9 and under), moderate-low (10-19, moderate (20-29), moderate-high, (30-39), high (over 39). For an explanation of the group dynamics structures, please refer to the previous section.

Investigating the teacher involvement, confusion, and group dynamics of each case further reveals the effectiveness of each team's communication. This summary is conceptually ordered to begin with the team that had the least confusion to the team that had the most confusion. Similar to the previous conceptually-ordered summary, the three game teams are at the top, while the three control teams are at the bottom. In this first section, each treatment will be summarized according to teacher involvement. In the second section, each case will be summarized according to their amount of confusion, how the team handled the confusion, the amount of off-topic conversations, when such conversations occurred, and whether teacher involvement was a factor. Findings will be supported with student quotations. Then, to conclude, a cross-case comparison will synthesize the differences in confusion, off-topic conversations, and teacher involvement based on treatment group.

Teacher involvement. As detailed in the above table, teachers were involved with the students differently in each treatment group. With the control group, teachers did a lot of whole class instruction and review, see Table 31.

Table 31

Visual and Auditory Evidence of Whole Class Involvement in the Control Group

Visual Evidence	Direct Instruction	Class Review
	Teacher A: A little professional tip. The hottest part of a flame is the very, very tip. So actually holding it, you don't have to go that far down. Just hold it right above, right there. Tilt it on its side and hold it right there. (*C1 Conversation, 858*)	Teacher A: I need a couple of volunteers as to how you would design an experiment. (*C1 Conversation, 143*)
	Teacher B: Guys, one test on the powder. Then record your results. (*C2 Conversation, 251*) Teacher B: Put some water in each of the powders so it becomes like a liquid. (*C3 Conversation, 466*)	Teacher B: Someone want to share their plan? (*C2 Conversation, 31*) Teacher B: Is there anything you noticed with the vinegar test, something that might have stood out with the vinegar test? (*C3 Conversation, 547*)

During the periods, teacher circulated around the room to help students in small groups, see Table 32. This interaction usually involved scaffolding some instructions for the students or asking questions to try and draw out their thinking a little bit further.

Table 32

Visual and Auditory Evidence of Teacher Scaffolding in the Control Group

Teacher A	Teacher Quotations	Teacher B
	Teacher A: Okay. Now what do we want to do next? (*C1 Conversation, 189*) Teacher B: What are you trying to do? What is the overall goal of this lab? (*C2 Conversation, 15*) Teacher B: Be specific. Why? What results? (*C3 Conversation, 1233*)	

Although not mentioned in the conceptually-ordered summary table (Table 29), teachers also maintained the role of classroom manager, making sure that everyone was physically staying with their team and getting their work done. Please refer to Table 33 below for evidence of managerial involvement.

Table 33

Visual and Auditory Evidence of Managerial Involvement in the Control Group

Visual Evidence	Auditory Evidence
	Teacher A: We are all working very well. Let's just keep the volume down a little bit. (*C1 Conversation, 545*) Teacher B: Why are you wandering around? Go back to your seat. (*C2 Conversation, 1328*) Teacher B: Where do you belong? Not with this group. (*C3 Conversation, 307*)

With the experimental group, teachers did not do a lot of whole class instruction and review, especially with the higher achieving teams. In fact, the dynamic between the teacher and team E1 was very different from the dynamic seen in the control group. The girls in E1 would check in briefly with the teacher when they needed an answer. For example:

S1: Do we put today's date?
T: That will suffice.
(*Team E1 Conversation, 74:75*)

In turn, the teacher would check in with the girls for a status from time to time:

T: Are you finished with your paper, ready to test your powder tomorrow?
S1: No, not yet.
T: Not yet, okay.
(*Team E1 Conversation, 683:685*)

In this team, there was no real evidence of instruction, scaffolding, or review.

Students in Team E2 needed some guidance with the technology and needed some corralling at the end of the periods. At the end of Day 1, Teacher B said to the class: "Ladies and gentlemen, I need you all to get all of your stuff and head back to the classroom right now. Make sure you hang onto the iPad. Don't drop them." (*Team E2 Conversation, 126*). While out in the hallways, teachers were available for assistance as needed, keeping a watchful eye over student progress but generally without instructing them, see Figure 15.

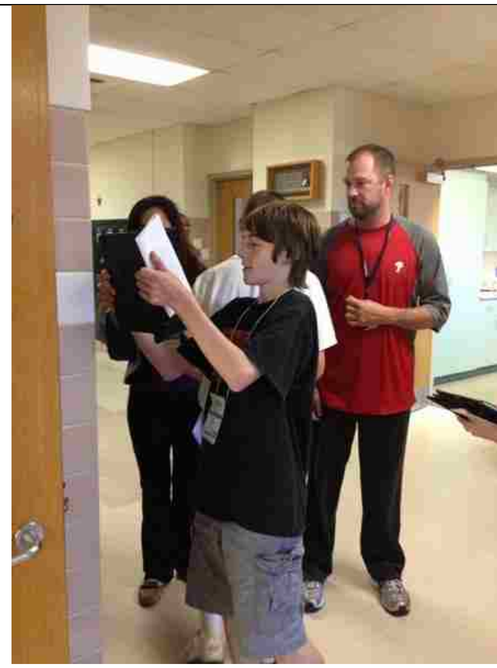
Teacher A	Teacher B
	

Figure 15: Examples of teachers' hallway interactions.

Overall, there was some evidence of direct instruction with comments such as "you should be writing down observations" (*Team E2 Conversation*, 502). But, more importantly, the teacher's role seemed to shift into one where the teacher was available to help when the students were ready. Especially in the room, teachers worked mostly with small groups as shown in Figure 16.

Teacher A	Teacher B
	

Figure 16: Teachers interactions in the classroom during the experiment.

Student teams would let the teacher know when they were ready for the next part of the activity. For example:

> S1: [Teacher's name], we are ready for the heat test.
> T: Awesome.
> *(Team E2 Conversation, 1242:1243)*

In the lowest achieving game team, the teacher did provide basic instruction and provided some scaffolding; however, this finding might be slightly skewed since the team had a dedicated aide for part of the intervention. Overall, teacher involvement in the treatment groups had significant differences. In the control, the teacher was mainly a sage of the stage; in the experiment, the teacher was involved mostly as a guide on the side.

Team E3: Expert guide. This team had a low level of confusion. During their moments of confusion, about half of the time someone was confused about what to write on the worksheet. Here are some examples:

> Should we put the secretary or not? (*Team E3 Conversation, 483*)
> Now we have to write what happened with the cornstarch. Okay, so, wait, I'm confused here. [*Teacher's name*]? I'm confused. (*Team E3 Conversation, 695*)

At other times, confusion stemmed from lack of knowledge on how to do something, such as: "how do you do the pH test?" (*Team E3 Conversation, 968*). Confusion was alleviated within the team or by teacher intervention. Sometimes, a team member took a second look or another member offered assistance. It could also be a combination of both means. For example:

> S1: I don't know which code to scan. Oh, scan this code.
> S2: You've got to go here.
> (*Team E3 Conversation, 10:11*)

At other times, the teacher stepped in to answer the question. Overall, the expert guide either helped team members solve issues within their team or—if he did not know—he deferred to the expertise of the classroom teacher. This team was also always on task and on-topic; they never talked off-topic except for one small comment about lunch.

Team E1: Community action. This team had a moderate to low level of confusion. Similar to the expert guide team, this team of girls exhibited confusion about what to write and how to proceed both within the game environment and during the hands-on lab. Due to the community action approach, when there was confusion over what to write, the girls often helped each. For example:

> S1: For the lead investigator, what do you write? Yourself?
> S2: Just write yourself.
> (*Team E1 Conversation, 71:72*)

Sometimes, the girl herself was able to alleviate her own confusion: "Wait, the steps to try, oh, to try to identify the mystery powder. Got it." (*Team E1 Conversation, 216*). During gameplay and during the lab, a team member was always there to lend assistance. Here are two examples:

> S1: Can we just rinse them out?
> S2: Yes. I think so.
> (*Team E1 Conversation, 575:576*)

> S1: How do you put this out? You just blow it out?
> S2: I don't know. Don't blow it out. Don't blow it out.
> S3: Maybe you just put the top on.
> S4: You just put the top on.
> (*Team E1 Conversation, 625:628*)

Overall, the girls consistently resolved their confusion with the community action approach, if they could not alleviate their own confusion, then a team member was there to help. They rarely—if ever—deferred to the teacher.

The girls had a moderate to low level of off-topic conversations. These conversations occurred at the very beginning of the period before the team had determined their first task or at the very end of the period as they were wrapping up the activity. The only other time when the girls steered off-topic was as they walked between game locations; no teacher involvement was necessary to get them back on task, but rather scanning the next code immediately brought their conversation back to the task at hand. There were no off-topic conversations during the hands-on experiment.

Team E2: Effective democracy. This team had a moderate level of confusion. Similar to the expert guide and the community action team, this team also showcased confusion about what to write and how to proceed both within the game and during the hands-on lab; however, they differed in how they resolved their confusion. Team members were there to help, but not always in a nurturing manner. For example, this interaction occurred early in the activity as they were trying to figure out how to use the game:

> S1: Wait, what do we do? Activate?
> S2: No, recorder, scan code, and then just put it right at the code. No, no, the code right there. Bam.
> (*Team E2 Conversation, 38:39*)

At other times, teammates were available to work out a solution with a teammate even if that solution included asking the teacher:

> S1: Does it say you unlocked chapter three? How come me and Peter don't have that?
> S2: I don't know. Maybe you should ask.
> …
> S2: Maybe you should ask her or him. He's over there.
> (*Team E2 Conversation, 862:866*)

Unlike the community action team, sometimes this team involved the teacher in order to stay on track and resolve their confusion over what to do next:

> S1: So how do we get the supplies? What do we do? Like do we just?
> T: You already tested the mystery powder?
> S1: Yeah.
> T: All right, let's rock and roll.
> (*Team E2 Conversation, 1075:1078*)

Some general confusion about the pH test persisted for the majority of the activity and it went unresolved for a long time. During the activity, one student referenced the pictures of the pH test that one boy had by saying "don't ask Terrell[4]. Terrell's is really confusing. We don't know yet." (*Team E2 Conversation, 678*). Several times, the team verbalized their trouble understanding the pH test and what observations to make. Eventually, they worked out together how to determine the pH level of the tested powder:

> S1: It turned green. Should I write that on the evidence table?
> …
> S1: It looks like a light green.
> S2: Try this one, pH 8. I think it's pH 8.
> S3: Um-hmm.
> S1: So write pH 8.
> (*Team E2 Conversation, 1179:1184*)

Overall, this team used different strategies to resolve their confusion; sometimes they worked well together, other times they helped each other in an unfriendly way, and they even deferred to the teacher in certain instances.

Team E2 had a moderate amount of off-topic conversations. Similar to team E1, this team went off-topic and the very beginning and the very end of period. Also similar to E1, this team had off-topic conversations while walking between game locations. In addition, some team members spoke off-topic while other teammates tried to determine what the next quest was or while they waited to scan a code. [Note: Teams had three to four members and only one person could scan a code at a time.] During the classroom experiment, this team had some brief off-topic conversations but never required teacher intervention to get back on track.

Team C3: Egalitarian leadership. This team had a moderate level of confusion. Some confusion stemmed from lack of understanding of the directions but it also derived from lack of content knowledge as well. The girl leader seemed to have the best grasp of both content and

[4] Boy's name changed to maintain anonymity.

instructions for the activity. When the boys were confused, the girl leader offered directional guidance. Sometimes she offered comforting reassurance; sometimes she was not as comforting. Here are some examples:

boy1: Then what do we do?
girl: Then we get the cornstarch, the sugar and the baking soda.
(*Team C3 Conversation, 59:60*)

boy2: I don't know what you're talking about. I don't understand what.
girl: I understand it.
(*Team C3 Conversation, 202:203*)

boy1: Drip it? Drip it where though?
girl: Wait, wait, wait.
boy1: On baking soda ….
girl: No, this one first.
boy1: How many?
girl: Do it again. Do it again.
(*Team C3 Conversation, 228:233*)

When she did not know how to alleviate confusion, she deferred to the teacher:

boy1: What do we do? What do we do?
boy2: I don't know.
girl: Excuse me, sir, we don't know what to do.
T: You have the instructions. What do they say to do?
girl: Oh, yeah, all right.
(*Team C3 Conversation, 435:439*)

Overall, this team used strategies to resolve their confusion that aligned to their group dynamics. The girl leader often resolved the confusion for her teammates, or when needed, asked for assistance from her teacher. There was no evidence of individuals resolved their own confusion.

Team C3 had a high level of off-topic conversations; off-topic conversations were consistently happening throughout the experiment with this team. Similar to other teams, team C3 went off-topic at the beginning and end of the period. This team talked off-topic at opportune moments such as when a teammate left the table to get the testing supplies. However, this team also spoke off-topic at inopportune times as well such as in the middle of an experiment. For example, here is a conversation about which high school the students were going to while trying to conduct the iodine test:

S1: Why do I want to go to Liberty. I might just go …. because I'm moving.
S2: You're going to ….?
S1: I might, but I kind of do want to go to Liberty. Look at the …. All right, so we've got to focus, iodine.
S2: I …. Liberty.
S1: What?
S2: You're going to Liberty.
(*Team C3 Conversation, 151:156*)

Additionally, this team had brief off-topic conversations when the teacher was giving instructions to the whole class as well as when they were supposed to be filling in the worksheet. At times, the teacher tried to get them back on task with comments such as "now you guys are filling this area out here" (Team C3 Conversation, 813). Overall, this team had a substantial amount of off-topic conversations throughout the entire activity.

Team C2: Dictatorship. This team had a moderate to high level of confusion. First, rather than resolving confusion by asking for teacher assistance, this team would instead started talking off-topic. Second, unlike game teams who responded to each other's confusion, questions and uncertainty would frequently go unresolved in this control team. Sometimes a teammate would ask a question, looking for clarification and the confusion was never alleviated and simply left unanswered. When one teammate asked "wait, which one did that?" no one answered (1208). This team also struggled to work together to alleviate confusion. Rather than having a team dialogue, one individual might offer a solution, or several students would offer their own solutions, with no real resolution. Here is an example:

S1: What do we do now?
S2: Now it's Brittany's turn to clean it up.
S3: Now we pick a mystery powder and test it like with all the tests.
S4: We have to write down all this other stuff.
(*Team C2 Conversation, 996:999*)

Rather than directing or reassuring his teammates in a manner similar to the egalitarian leader, the dictator simply responded to questions of uncertainty with "just read my paper" (*Team C2 Conversation, 328*). Overall, this team resolved their confusion in ways that aligned to their group dynamics. As the dictator, the boy might determine the next direction without any team input or, when there was team input, the best course of action was not openly discussed. The dictator did not even defer to assistance from his teacher. Just as in the egalitarian leadership team, there was no evidence of individuals resolving their own confusion, if anything, confusion was left unresolved.

Team C2 had a moderate to high level of off-topic conversations. Besides the beginning and end of class, and when teammates left the table to collect supplies, this team spoke off-topic the most when they were supposed to be filling in the worksheet. The teacher would try to get them on track with comments such as "think about the results you got from doing this sugar and compare it to doing these two sugars" (*Team C2 Conversation, 1379*). Unlike C3, team C2 did not generally talk off-topic when the teacher was giving instructions or during their experiments. However, towards the end of the activity when tension among the team was at its highest, some brief off-topic dialogue appeared during experimentation.

Team C1: Ineffective democracy. This team had a moderate to high level of confusion. More than any other team, a significant portion of their confusion was over which powder they were testing. Here are some individual examples:

> Wait, which one is this? (*Team E1 Conversation, 328*)
> The sugar is this; wait, which one's the sugar? (*Team E1 Conversation, 338*)
> Wait, which is cornstarch? (*Team E1 Conversation, 381*)
> What's this one? Baking soda? (*Team E1 Conversation, 455*)

Such questions would generally promote a short team discussion, but the dialogue often revealed even more confusion:

S1: You’re on the wrong one.
S2: That’s cornstarch.
S1: That’s cornstarch.
S3: Aren’t we on baking soda?
S1: No. It’s sugar. This one’s sugar.
(*Team C1 Conversation, 201:205*)

Beyond uncertainty about their powders, the team also had general uncertainty that was not alleviated. One teammate asked, “How are we going to do this?” and was never answered (*Team C1 Conversation, 68*). When there was confusion over what to write or what direction to take, someone on the team often addressed confusion with a direction. But just giving the command was not enough to alleviate the confusion and the conversation sometimes drifted off-topic at that point:

S1: Okay, and then what?
S2: You just get it a little wet, and then you put this on, and then put the sugar.
S1: So you put water in this ….? …. …. And then you put that on.
S3: These things taste weird.
S2: How do you know?
S3: [Teacher’s name] made us all taste ….
(*Team C1 Conversation, 427:432*)

This team had a type of group interaction that gave the illusion of having a solution but there was still underlying uncertainty. Ultimately, with this type of ineffective democracy, confusion was a persistent problem; they very rarely looked to the teacher for guidance and they were unable to communally resolve their own confusion.

Team C1 had a moderate level of off-topic conversations. Beyond the typical off-topic conversations at the beginning and end of the period, this team had brief off-topic conversations most often during the writing portions of the class periods. Team members were easily distracted by the recording device as well as their pencils, erasers, the way each other wrote certain letters and more. However, these conversations were typically brief and resolved on their own without teacher intervention.

Summary of findings. This summary will synthesize the findings from team confusion, off-topic conversations, and teacher involvement. In conclusion, the game teams managed their confusion more effectively through group dynamics resolving their confusion as individuals or as a collective. They used their teachers for instructional support when necessary; in turn, the teacher shifted away from the traditional teacher in front of the classroom to more of a guide on the side. Despite less direct teacher involvement, game teams actually talked off topic less that control teams. When game teams talked off-topic, it generally occurred at opportune times that did not impede the learning process, such as the beginning and end of the period.

In contrast, control teams did not resolve their confusion through discussion. In cases where team dialogue existed, it only temporarily masked the confusion rather than fully resolving it. In other cases, team leaders just determined the solution and executed it without buy-in from the team. There was no evidence of individuals resolving their own confusion in control teams. Also, control teams did not generally rely on the teacher for assistance but rather the teacher would get involved when the teams went off-task. In fact, control teams talked off-topic more than game teams. They spoke off-topic during opportune times as well as inopportune times. Teams in the control drifted off-topic when they were working on the worksheet and even during experimentation. In one extreme case, a team was talking off-topic during teacher instruction. Control teams required some teacher intervention to stay on-task; game teams had less teacher involvement yet remained on-topic more and had less confusion overall.

CHAPTER 5: DISCUSSION

The previous chapter presented the findings for the five research questions investigated by this study. This study is structured around three themes: flow, scientific practices, and collaboration. In this chapter, I bring together the quantitative and qualitative findings for each theme explaining how the findings connect to the literature and discussing why they occurred. In addition, I specifically address the relationship between flow and learning which aligns with all three themes. I conclude by synthesizing all of the themes into a theoretical proposition for how and why collaborative learning occurs in INPLACE mobile games.

Connections to Literature on Learning Science through Games

Prior research indicated that collaborative mobile games held promise for promoting scientific practices during gameplay. Yet, Young et al. (2012) reported that there was limited evidence to support video games increasing student achievement in K-12 education. This study set out to document student achievement in terms of scientific practices in two groups: game players and control participants who used business-as-usual classroom laboratory learning. In this study, culminating artifacts created by all students in the study represented some level of each of these scientific practices: defining the problem, planning the investigation, discussing data, constructing explanations, arguing with evidence, and communicating scientific information. Statistical analysis showed that these artifacts, or written reports, created by game players showed *significantly* higher levels of scientific practices than control participants. Results from the Apple Classrooms of Tomorrow (ACOT) project supported this finding; during the ACOT project students who used technology effectively in groups produced higher quality artifacts (Fisher et al., 1996). In addition, findings from Ketelhut et al. (2010) further reinforced this finding; students working in collaborative groups within the *River City MUVE* produced written letters that scored more than double that of their control counterparts on measures of scientific inquiry.

Empirical Evidence

One of the criticisms against game-based learning is lack of sufficient empirical evidence demonstrating that games provided more knowledge gains than other instructional modes (Young et al., 2012). Annetta et al. (2009) found no statistical difference between post-test scores of 66 game participants who played a teacher-created genetics video game and 63 control participants who received traditional instruction. In addition, Wrzesien and Alcañiz Raya (2010) compared traditional ecology instruction to participation in an immersive virtual simulation and did not report any statistical differences in learning outcomes between treatments groups. Unlike these earlier studies that found no statistical significance between treatment groups, the study results reported in this dissertation found significant statistical difference between the report scores of the game players and control participants who received business-as-usual instruction.

Experts have reported concern for measuring game-based learning with traditional assessments; therefore, knowledge gains were assessed differently in this study from the previous studies in which no statistical differences were found. This study required students to write reports since research showed that open-ended explanations captured student's knowledge integration ability better than multiple-choice items (Hee-Sun et al., 2011). Wrzesien and Alcañiz Raya (2010) and Annetta et al. (2009) both used traditional content post-tests with multiple choice items; this study used only open-ended explanations—there were no multiple-choice items. The open-ended explanations afforded a more authentic assessment: writing an incident form. Ketelhut et al. (2010) compared an authentic assessment where students wrote a letter to the mayor to a traditional multiple-choice test. On the authentic assessment, students who participated in guided social constructivist learning within the River City multi-user virtual environment scored higher that control students on all science inquiry tasks except stating the problem; however their multiple-choice test scores were not significantly different than the control. Ketelhut et al. (2010) concluded that the standardized content test does not accurately measure the scientific practices gained during gameplay; game-based learning is best measured with an authentic assessment. Based on recommendations from previous research it was important to use an open-end assessment that was graded with a rubric and this study found statistical differences on the scientific practices between treatment groups.

Case Study Evidence

This study not only offered empirical support for significant knowledge gains of the game players over control participants but also the ways in which game teams differed qualitatively from control teams. In brief, game teams showcased a deeper ability to integrate the scientific practice into their conversations than control teams. Specifically, game teams offered observations that were more specific and substantive than control teams. Game teams also revealed a stronger understanding of describing the problem as well as some understanding of how to create a hypothesis. In terms of explanations, below average students struggled with this practice in both treatments; however, average and above average game teams constructed more explanations than their control counterparts. All game teams argued with evidence; not all control teams engaged in this practice. The only scientific practice in which control teams showcased a better understanding than game teams was detailing their plan for determining the components of the mystery powder.

An important finding from this embedded qualitative analysis was the strong evidence of scientific inquiry that players revealed during gameplay. Evidence of scientific inquiry during game play has been found in earlier qualitative studies (Asbell-Clarke et al., 2012; Squire & Jan, 2007). Previously, Squire (2010) found that groups playing *Sick at South Beach* argued quite intensely for their role's scientific point of view. Also, Squire and Jan (2007) explored scientific argumentation using a case study comparison of three different groups playing *Mad City Mystery* and found that all groups engaged in argumentation cycles. Specifically, Squire and Jan (2007) reported that during the game, players weighed evidence, developed hypotheses, tested them against evidence, and generated evidence-based theories. In this study, game teams defined the problem, discussed observations, analyzed and synthesized data, constructed explanations, and argued with evidence. Unlike Squire and Jan (2007) who were not able to triangulate their data with quantitative results, this study has statistical data to support broader assertions about what players learned. This study not only qualitatively showed evidence of more scientific practice within game teams but also statistically reported higher levels of scientific practice as represented in their written reports. The qualitative findings from this study build on the work of Squire and Jan (2007) and Squire (2010) by comparing three game teams with three control teams and triangulating the findings with statistics.

In this study, multiple members of game teams were engaged in evidence-based argumentation, hence making that scientific practice highly collaborative; Asbell-Clarke et al. (2012) found similar results. Asbell-Clarke et al. (2012) reported that *Martian Boneyards* nurtured sustained scientific inquiry during game play. They found that players used language suggestive of evidence-based argumentation and that inquiry among players was highly collaborative. In the current study, only one control team had a conversation in which multiple team members argued with evidence; this was a notable contrast to game teams. Essentially, control teams did not have the same level of argumentation as game teams nor were such conversations collaborative. Ultimately, the findings from this case study are similar to the conclusion of Cheng and Tsai (2012) who reported that location-based AR games that incorporate collaborative role-playing design encourage players to observe scientific phenomena, ask questions, investigate data, create hypotheses, and construct explanations.

Related to these findings on scientific conversations, this study also determined that game teams had fewer occurrences of off-topic conversations while control teams had more occurrences. Asbell-Clarke et al. (2012) reported a similar finding that non-game related discussion was notably low.

Anecdotal Evidence

The last piece of evidence that revealed the capacity of game teams to argue with scientific evidence is best described by this anecdote. Unfortunately, in one of the treatment classrooms, the researchers mislabeled the bins containing the mystery powders. Depending on the game version being played, students were told to take mystery powder A or mystery powder B. *Correctly* analyzing the *wrong* powder meant that the detective in the game told students that they were wrong and asked them to re-analyze the results. Mislabeling the powders forced students into this situation. Game teams in this classroom were so convinced that their lab results were correct that they argued with the teacher using evidence from their experiment to support their claim that the game had to be wrong. It was the conviction of the students that led the researcher to test the powders, which confirmed that indeed the powders were improperly labeled and the students had analyzed their data accurately. Similarly, Barab et al. (2009) reported that student teams who experienced scientific practice in an immersive game performed better on a performance-based transfer task than control teams. In this study, game teams performed extremely well on the mystery powder experiment which required them to integrate knowledge that they had learned in the game.

INPLACE Mobile Games Lead to Deeper Scientific Practices

This study has confirmed and extended the findings of previous research studies on game-based learning in science education. In this section, I will discuss my rationale for why game teams in this study had more scientific practice in their reports and conversations. Generally, INPLACE mobile games support students' newer, neomillenial learning styles which consist of preferences for learning environments that support communal learning, active learning, and nonlinear learning (Dieterle, 2009). Specifically, INPLACE mobile games lead to deeper scientific practice by supporting situated learning, social construction of knowledge, and autonomy.

Meaningful learning through situated learning. As mentioned in Chapter 2, mobile devices afford portability making them effective tools for situated learning (Dede, 2009; Fotouhi-Ghazvini, et al., 2009; Rosenbaum et al., 2007). While it could be argued that students in the control were learning by doing—the colloquial term for situated cognition —it is the element of *lived practice* that distinguishes situated learning (Lave & Wenger, 1991). In the game teams, students experienced lived practice and it offered a more robust template upon which game players applied their learning. Control teams did not experience lived practice, and missed out on a powerful learning opportunity.

To experience situated learning, students were learning in the real world. Research conducted using *Quest Atlantis* concluded that the game-based curriculum's effectiveness derived from the notion that "it is more consistent with how the content is experienced in the real world" (Barab et al., 2010, p. 405). In this study, the game was designed in a similar fashion;

content was experienced in the real world of school. QR codes were placed in relevant locations around the school making the entire school into a crime scene complete with suspects, evidence, and mysterious substances. Essentially, the game immersed students in the lived practice of being a crime scene detective. Anecdotally, I witnessed a contrasting situation where a subset of students was robbed of the lived practice element of the game and it negatively impacted their experience. As mentioned in Chapter 3, a subset of students was dropped from the sample because a walkathon interfered with class time and all QR codes were moved into the classroom. Without the context of the school environment, this subset of students lost the opportunity for lived practice, and many students were visibly disappointed. One student specifically told me that the activity really lost something when you do not get to go in the hallway; I concurred. Students not only vocalized their disappointment but also reported very low flow experiences scores.

Meaningful learning through social construction. According to Reiser et al. (2012), students need to "actively listen and respond to one another" in order to be engaged in meaningful scientific practice (p. 36). Game teams had more engaged communications responses along with higher levels of communal language and lower levels of commands. These communication patterns seemed to connect with better group dynamics and more effective team communication. In brief, game teams met the precursor for meaningful learning by communicating well. Unfortunately, control teams had higher rejecting responses, higher commands, and less communal language. This pattern seemed to connect with less effective group dynamics and poor team communication skills. Control teams did not meet the precursor for meaningful learning since they did not communicate well.

In this dissertation study, game teams communicated better because their roles were defined within the game and not everyone was privy to same information; control teams struggled to understand their individual roles within the group and their group dynamics suffered. Previous games have served to create better communication among players by designing guidance into the game. Research has showed that when students were guided to socially construct their knowledge in the *River City MUVE*, they had a stronger understanding of scientific inquiry than other students (Ketelhut et al., 2010). Specifically, students in the guided social constructivist category showed a stronger understanding of scientific inquiry than not only control students but also the expert modeling-coaching and legitimate peripheral participation categories of *River City* players.

In terms of scientific communication, control teams did not have the same level of argumentation as game teams. Berland and Reiser (2009) posited that for students to truly engage in meaningful scientific argumentation and explanation, they need to be arguing for their own ideas, not because their teacher asked them to do it. Throughout the game, argumentation appeared multiple times with different players engaging in the practice. Players were discussing their own ideas as a group and struggling to come to group consensus. The game seemed to give them enough fodder for discussion and direction while allowing players to interpret and act autonomously according to their own instincts.

Meaningful learning through autonomy. According to Bruner (1961), a learner should become “as autonomous and self-propelled a thinker” as possible (p. 23). Game teams were fairly autonomous; they used their teachers for support when necessary. Control teams were not generally autonomous; they received direct instruction and—when they did work on their own—they did not rely on the teacher for assistance but rather the teacher would get involved when the groups went off-task. Students on the game teams had less teacher involvement yet better learning outcomes.

In this study, more autonomy and less direct instruction made for a meaningful learning environment. This finding supports the research of Bonawitz et al. (2011); the researchers studied young children playing with a new toy and reported that direct instruction limited spontaneous exploration and discovery. Preschoolers who were given instructions about the function of the toy spent less time playing with the toy and discovered less additional functionality of the toy than children who were not given direct instruction. Bonawitz et al. (2011) concluded that when kids are curious, they actually want to learn more.

The choice to limit direct instruction for game teams and switch to learning facilitation was not a requirement for teachers in the study. In fact, teachers were not told to behave in any specific way; they naturally taught in a way that made sense to them given their materials and the needs of the students. As Asbell-Clarke et al. (2012) found in their game *Martian Boneyards*, a particular character was crucial for encouraging reflective thoughts and facilitating learners thinking. In this study, the teacher "played" that role in the game. Just as a virtual character naturally infused conversations during *Martian Boneyards* with questions to prompt deeper inquiry, the actual teachers in this study could guide game teams in a similar fashion.

The nature of the game play meant that learners had autonomy to discover the information for themselves and seek guidance when and if needed. In the game, students were more autonomous; teachers were no longer delivering content but supporting students in their endeavors to discover it themselves. As Richardson (2012) explained, the focus of our educational system needs to shift from content mastery to learning mastery. Game players were experiencing learning mastery. They took more ownership over their learning, used their teachers when needed, and resolved confusion on their own or in conjunction with their teammates. Control teams were stuck in content mastery. Teachers provided direct instruction and had to keep them on task and on track.

In game teams, teachers had less direct involvement thus students were more intrinsically motivated; they took initiative and felt in control. Piaget (1962) believed that a child's insatiable curiosity drives their learning; furthermore, while playing, a child can initiate and control the activity. In game teams, it seemed students had more of an internal drive to learn rather than an external force directing them to learn. Ultimately, during play, children assimilated their learning (Piaget, 1962). If game players were truly experiencing this internal drive to keep playing, they could better assimilate scientific information and write reports with more scientific practice.

In summary, artifacts created by game teams showed higher levels of scientific practice and their conversations showed more evidence of scientific practice than control teams. This result is best explained by game teams experiencing more meaningful learning through their autonomous discovery, situated cognition, and social construction of knowledge. The next section will discuss the findings about flow experience among student teams and why the game teams had higher flow than the control teams.

Connections to Literature on Promoting Flow in Game-Based Science Learning

Prior research indicated that collaborative mobile games held promise for engaging students by promoting flow during gameplay. Qualitatively, researchers have confirmed improved student engagement while playing mobile learning science games including *Alien Contact* (Dunleavey et al., 2009) and *Power Agent* (Gustafsson et al., 2009). Researchers have also confirmed that participants' level of engagement had a statistically significant increase when playing a SEG (Annetta et al., 2009). In this study, engagement was measured according to flow theory; self-report surveys from game teams represented significantly higher levels of flow than control teams.

Previous research supported this finding. First, Park et al. (2010) used a self-report survey to measure flow of participants in a paper-based physical exploration, mobile-based physical exploration, and a virtual- game-based simulation with no physical exploration. The survey items captured player’s feelings pertaining to learning control, attention focus, cognitive curiosity, and intrinsic interest. In all cases, except learning control, players using the mobile-based version reported the highest levels of flow. Second, Ryu and Parsons (2012) used a similar self-report survey to measure cognitive curiosity, intrinsic interest, and risk-taking—the three categories associated with flow that the researchers wanted to examine. They compared reported flow levels from those participating in a collaborative mobile activity and a solitary mobile activity; collaborative mobile participants had higher flow scores in all categories than those in the solitary scenario. Unlike Ryu and Parsons (2012), who compared flow among collaborative and solo mobile participants, the current study compared a collaborative mobile activity with a collaborative non-mobile activity. Ryu and Parsons (2012) kept the variable of mobility constant; while this dissertation kept the variable of collaboration constant. Both studies found that participants in the collaborative mobile scenario reported the highest levels of flow. Lastly, game teams reported not only higher levels of flow but also had higher report scores than control teams. Admiraal et al. (2011) reported a similar finding; team flow seemed to be connected to how well the group performed in the game.

INPLACE Mobile Games Promote More Substantive Flow-Like Experiences

This study has confirmed and extended the findings of previous research studies promoting flow in collaborative mobile activities. In this section, I will discuss my rationale for why game teams in this study reported higher levels of flow than control teams. First, game characteristics were similar to the game studied by Bressler and Bodzin (2013); they found the game design seemed to promote the right challenge-skill balance for players. In fact, Bressler and Bodzin (2013) stated that challenge-skill balance was one of the most highly encountered flow elements along with total concentration and an intrinsically rewarding experience. In this study, students playing the INPLACE mobile game not only encountered those three flow elements but also clear goals and sense of control.

Challenge-skill balance. Game teams seemed to experience a better challenge-skill balance than control teams. According to Csikszentmihalyi (1997), “flow tends to occur when a person’s skills are fully involved in overcoming a challenge that is just about manageable” (p. 30); in other words, there is a good balance between the challenge presented and the person’s skills. For game teams, the challenge offered by the activity and the skill set of the group seemed to be in balance. When groups were challenged or confused, they managed their confusion effectively through group dynamics resolving their confusion as individuals or as a collective. If they could not overcome the issue through group discussion and they determined that teacher assistance was needed, then they would ask for help. Unfortunately, when control teams were challenged, they did not resolve their confusion as well. Control teams did not resolve their confusion through group dialogue. In cases where group dialogue existed, it only temporarily masked the confusion rather than fully resolving it. In other cases, group leaders just determined the solution and executed it without buy-in from the group. There was no evidence of individuals resolving their own confusion in control teams. Ultimately, the challenge of the activity was not in balance with the skills of the control teams, and they experienced lower levels of flow.

Sense of control. Game teams seemed to have a stronger sense of control than teams in the business-as-usual activity. In the experimental group, teachers did not provide a lot of direct instruction, but rather game teams used their teachers for instructional support when necessary. This dynamic occurred naturally; the teacher shifted away from the traditional teacher in front of the classroom to more of a "guide on the side". Lower levels of direct instruction seemed to give the game teams a sense of control. The force that propelled students through the game was more intrinsic rather than external. According to Csikszentmihalyi et al. (1993), people who are in flow are driven by intrinsic motivation.

On the other hand, teachers provided a lot of direct instruction and reviewing for the control group. Control teams received a blend of traditional teaching from the front of the room with some small group guidance during the lab experiments. Higher levels of teacher involvement gave control teams a lower sense of control. The force that propelled students through the activity was generally external—it was the teacher. Control teams spoke off-topic more frequently and required teacher intervention to remain on-task. Overall, control teams had a lower sense of control than game teams which contributed to lower levels of flow.

Total concentration on the task. Game teams seemed to be more concentrated on the task than control teams. In general, control teams spoke off-topic more than game teams. Control teams drifted off-topic when they were working on the worksheet and even during experimentation. They even required teacher involvement to stay on-task. In contrast, game teams required very little teacher invention to stay on task and generally talked off-topic only at opportune times that did not impede the learning process. Staying on topic meant they were concentrating more deeply on the learning task. It seems game teams achieved *the intellectual ideal*, a balance of playfulness and seriousness (Dewey, 1910). They enjoyed playful off-topic conversations while walking between QR codes, yet engaged in serious conversations once they received the next set of game information. Taking brief mental breaks between game elements may make it easier to concentrate more deeply when on-task.

Clear goals. Game teams seemed to understand their goals more clearly than control teams. According to Csikszentmihalyi (1997), “flow tends to occur when a person faces a clear set of goals that require appropriate responses” (p. 29). In the game, tasks were broken into quests making each step a clearly delineated goal. When a quest was completed, it would disappear from view so that a player’s screen emphasized only quests yet to be done. In addition, quests were divided evenly between all players giving each member a unique set of goals. All of these factors afforded not only clear goals but also seemed to encourage more orderly group dynamics. As discussed in Chapter 4, the three game teams were termed *expert guide, community action,* and *effective democracy*.

In the control activity, there was a handout with a long list of steps to follow; however, it was unclear how the group should organize themselves in order to accomplish their tasks. When no order was imposed by the activity, groups instituted their own order. In one control team, disorder took precedence; they were the *ineffective democracy*. In another team, order was imposed by a group member and it offered some benefits; this was *egalitarian leadership*. In the last team, order was imposed by one member yet it served as a destructive force in the team; this was the *dictatorship*. All in all, game teams experienced clear goals individually and collectively. Interestingly, Huizinga (1938/1971) made the case that one of the most positive features of play is that it creates order, it has rules. The game teams had more order, and that led to more enjoyable, playful experience.

Intrinsically rewarding experience. Research on students using technology in groups has shown high levels of enjoyment. During the Apple Classrooms of Tomorrow project, students working in groups demonstrated intensified engagement (Fisher et al., 1996). Folta (2010) concluded from her research on a mobile AR science game that “the community atmosphere appeared to be an important part of the participants’ enjoyment and learning experience” (p. 127). In the original version of *School Scene Investigators*, students particularly enjoyed working together (Bressler & Bodzin, 2013).

In this study, game teams seemed to intrinsically enjoy the experience more than control teams. General enthusiasm was observed on many occasions; however, there were even some student statements recorded. At the beginning of the game, one girl said, “It seems so fun. I’m so excited.” At the end of the game, one boy said, “This was actually really fun.” Enjoyment among game teams may have been partially derived from their successful communication patterns. Game teams produced a fairly high level of engaged responses. They also had higher levels of communal language and lower levels of commands. In contrast, control teams produced a fairly high level of non-engaged responses. They had lower levels of communal language and higher levels of commands. In general, these communication patterns could be seen as a source of enjoyment for game teams, and possibly a source of frustration for control teams. When teammates support your ideas or are at least willing to discuss them, it is more enjoyable than when teammates reject your ideas. Such enjoyment would contribute to higher flow scores for game teams, while frustration among control teams would explain their lower flow scores.

In summary, game teams reported higher levels of flow than control teams. This result is best explained by game teams experiencing these specific elements of flow: challenge-skill balance, sense of control, total concentration on the task, clear goals, and an intrinsically rewarding experience. The next section will discuss the findings about the relationship between students' flow experiences and their learning outcomes.

Connections to Literature on the Link between Flow and Learning

Bressler and Bodzin (2013) indicated that there is a role for INPLACE mobile games in learning; specifically, in the original *School Scene Investigators*, players persevered at challenges, and upon success, desired greater challenges. Yet, no prior research has shown empirical data to prove that flow increases learning in an INPLACE mobile game. This study specifically examined the empirical connection between flow and learning. A multiple analysis of variance was utilized so that the variables of flow score and report score (learning outcome) could be investigated in the same statistical model. The MANCOVA results showed multivariate significance; therefore, some relationship existed between the dependent variables.

Since research indicated that flow could lead to improved learning, this study tested for a predictive relationship of flow on learning. A multiple hierarchical regression was used in order to control for certain variables before testing the relationship of flow on learning as quantified by report score. Math achievement was entered in block 1 as a control variable; it had a significant predictive relationship on report score which was expected. Treatment group was also entered in block 1 because the study was interested in testing the predictive relationship of flow on learning regardless of treatment group. Flow score was the only variable included in block 2. Notably, treatment group was significant in the first model, but it was no longer significant in the second model. In Model 1, treatment group represented the distinction of game teams having higher flow scores and control teams having lower flow scores; however, in Model 2, treatment group is no longer significant because flow score is present and it represents the real reason why treatment group was significant in the first model. There is a small predictive relationship between flow score and report score; regardless of treatment group, a higher flow score predicted a higher report score. This means that learning experiences that promote higher flow levels should be considered as desirable learning experiences for all types of students whether that is an INPLACE mobile game or a hands-on collaborative experiment.

Flow experience, as reported by each individual student, uniquely predicted approximately 3% of the variance in his or her report score which was found to be statistically significant. This result stands in contrast to previous research results. As discussed above, several studies have indicated higher levels of flow in collaborative mobile activities, but no research had revealed a connected between flow and learning. Admiraal et al. (2011) reported that student learning outcomes from the mobile AR history game called *Frequency 1550* were not related to flow during gameplay. Additionally, Park et al. (2010) found no difference between the learning benefits of a physical exploration played on a mobile device versus one played with pencil and paper, despite the fact that the mobile device promoted higher flow.

This study may have uncovered a connection between flow and learning because previous studies did not measure all elements of flow, their learning experiences were of shorter duration, and their learning measures did not completely capture all the learning benefits of the mobile activity. First, this study measured flow experience with a self-report survey that captured all nine elements of flow as outlined by Csikszentmihalyi (1996). Park et al. (2010) only measured a few elements of flow including, control, concentration, cognitive curiosity, and intrinsic interest. They did not assess challenge-skill balance— an important factor of flow for INPLACE mobile games according to Bressler and Bodzin (2013)—or clear goals. Admiraal et al. (2011) did not use a self-report survey but instead relied on observations of student's experiencing flow which they acknowledged did not capture the subjective experience of flow.

Second, the duration of the learning experience for this study spanned four to five days, during one class period on each day. In contrast, Park et al. (2010) and Admiraal et al. (2011) conducted their interventions on only one day. Finally, this study measured learning outcome with a culminating artifact that included all open-ended questions. Previous researchers either used a knowledge test with a blend of multiple choice and open-ended questions (Admiraal et al., 2011) or simple content recall (Park et al., 2010). Future research that investigates the relationship between flow and learning should measure all elements of flow, conduct interventions across multiple days, and measure learning outcomes with an authentic open-ended assessment.

(Social) Flow and (Social) Learning

Flow experience had a statistically significant predictive relationship with learning, but it only predicted 3% of the variance in report score. The entire model of math achievement, treatment group, and flow score only predicted about 16% of the variance in report score. Obviously, some variables were absent from the model. Most notably, the model does not account for the social aspects of learning. In this case, the social interactions within teams played a sizable role in group success, documented in the communication patterns and group dynamics.

Social factors have a complex relationship with the model because social factors can individually impact flow (the predictor) as well as learning (the outcome), especially in game-based learning. At the Boston Museum of Science, Klopfer et al. (2005) found that the INPLACE mobile game *Mystery at the Museum* fostered feelings of connection between the visitors. Feedback from players confirmed that they were engaged with their fellow players. A positive connection to teammates can even deepen player engagement with the game. Researchers have found that player socialization is connected to their engagement with the game (Barendregt & Bekker, 2011). While some aspects of such engagement are attributable to an individual's flow, there may be another factor involved called social flow. According to Walker (2010), "flow in a social context may be a qualitatively different phenomenon than flow experienced in isolation" (p. 4). Specifically, he found that experiencing flow as a group is more enjoyable than experiencing it alone.

Walker (2010) investigated social flow and found that people who collaborated interdependently experienced more enjoyment in flow than those who worked with less interdependence. Interdependency is a popular strategy for promoting collaboration in mobile AR games. Several researchers in the field of mobile learning games incorporated this concept into their game designs by creating interdependent roles (Bressler & Bodzin, 2013; Dunleavy, et al., 2007; Dunleavy & Simmons, 2011; Klopfer et al., 2005). Bressler and Bodzin (2013) concluded that the interdependent nature of gameplay in the original *School Scene Investigators* seemed to increase social flow.

Social interdependence is also a way to promote effective group learning. Aronson and Patnoe (2011) explained why interdependence improves learning for those present in the group:

> Learning from each other gradually diminishes the need to try to outperform each other because one student's learning enhances the performance of the other students instead of inhibiting it, as is usually the case in most competitive, teacher-oriented classrooms. (p. 10).

Driscoll (2005) summarized Vygotsky's requirement that social interaction must have intersubjectivity to promote effective learning, explaining that students must "co-construct the solution to a problem or share in joint decision making" (p. 258). Collaborative game-based learning scaffolds such interaction using the jigsaw technique (Aronson, 1978; Aronson and Patnoe, 2011). Demetriadis et al. (2012) posited that the sociocultural learning that takes place within games works best when there is shared power and authority through scripted collaboration. Therefore, this study went beyond jigsawing information and divided up quests evenly between all players to encourage shared leadership of the group making each member a critical component of the whole.

Given the complicated relationship between social interaction and flow along with social interaction and learning, social interaction could be a moderating variable in the predictive relationship of flow and learning. More research is required to explore the empirical relationship of social interaction and how it moderates the connection between flow and learning. The next section will discuss the findings about collaboration among student teams and shed light on the association between game mechanics, social interactions, and collaborative learning.

Connections to Literature on Mobile Games and Collaboration

Prior research indicated that collaborative mobile games held promise for promoting effective collaborative practice by scaffolding and supporting discourse during gameplay. The literature review called for more investigations examining how game mechanics and social interactions contribute to learning and achievement gains (Sharritt, 2008; Young et al., 2012). Specifically, Clark (2007) claimed that future research needed to make a better attempt at valid comparison groups. In this study, student teams from both the experimental game and the control activity were analyzed in a cross-case comparison that emphasized their conversational dialogue.

As discussed above, game teams score significantly higher on their reports and reported higher levels of flow. Additionally, in the cross-case study comparison, game teams revealed a deeper ability to integrate most scientific practices into their conversations. Thus, generally speaking, game teams were more successful because they demonstrated stronger scientific practices and higher report scores. Communication responses in the game teams and the control teams revealed specific patterns that indicated their likelihood for success.

Barron (2003) investigated how problem-solving outcomes were influenced by group interactions. According to Barron (2003), successful groups produced a fairly high level of engaged responses—accept and discuss responses—in comparison to their non-engaged, reject responses. In this study, games teams had this pattern of responses. According to Barron (2003), less successful groups produce a fairly high level of non-engaged responses in comparison to their engaged responses. In this study, control teams had this pattern of responses. The results of this study are consistent with Barron's (2003) findings: control teams had less successful outcomes and used more non-engaged responses while game teams had more successful outcomes and used more engaged responses.

Going beyond Barron's (2003) work, this study discovered another communication difference between the more successful game teams and the less successful control teams: the use of communal language versus command language. Game teams used not only more engaged responses but also higher levels of communal language. They frequently addressed the group collectively and often referred to the group as an entity with words such as "we," "we're," and "let's." In Team E1, their communal language served not only to promote group cohesion but also to keep them on track towards their goals. Similarly, Team E2 referred to their team as a group entity with specific goals to achieve. Teams E1 and E2 asked each other a lot of questions. For Team E3, they had ongoing conversations to develop their communal understanding of the constraints of the main problem posed by the game. Group communication patterns exhibited by game teams seemed to connect with better group dynamics and more effective team communication. Unfortunately, control teams showcased communication patterns that seemed to connect with less effective group dynamics and less effective team communication. Control teams used not only non-engaged responses but also higher levels of command language; they were frequently telling each other what to do by using directive language such as "go get the other one." Overall, game teams used more communal language and more engaged responses and were more successful collaborators while control teams used more command language and more non-engaged responses and were less successful collaborators.

INPLACE Mobile Games Promote Successful Communication and Collaboration

In terms of design, the big difference between treatment groups was that the game environment provided more collaborative scaffolding than the control activity. Games teams were more successful because the interdependent roles that were designed into the game promoted positive communication patterns and a shared leadership structure; unfortunately, without such scaffolding to facilitate productive communication and collaboration, control teams were less successful. This supports the findings of Demetriadis et al. (2012) who summarized the relevant research on freely collaborating students and concluded that productive learning interactions do not occur when collaborative groups have no support or scaffolding. Not only do students need support, but the scaffolding needs to promote interdependency (Vygotsky, 1978). In this study, game teams were able to share decision making and co-construct solutions because of the interdependency of their collaboration as designed into the game mechanics, which in turned facilitated better communication and positive social interactions.

Game mechanics. Squire and Jan (2007) argued that the success of *Mad City Mystery* was achieved through game mechanics; roles and collaboration were two facets of gameplay that scaffolded student thinking. In this section, I will add to this conclusion by explaining how roles also contribute to player satisfaction and the overall success of the collaborative group.

Roles. To summarize the collective viewpoint of the play theorists from Chapter 2, play is directed, it has order, and it has rules. Having players select roles provides game designers with a mechanism for establishing some of the social rules of gameplay. In describing the rationale behind game roles, Salen and Zimmerman (2004) claimed that “we don’t mean that a player becomes a character in a story…we mean that each player has a role in the social network of a game” (p. 463).

Previous research has shown that such game roles are not only tied to successful collaborations but also scaffold scientific thinking and improve player satisfaction; these conclusions support findings from this study. First, game teams showed higher levels of scientific practice in their conversations and reports in part because they played roles, whereas the control teams did not. Similarly, Squire and Jan (2007) found that the game roles gave students the chance to see themselves as investigators—rather than students—and this helped to scaffold their scientific thinking. Second, game teams showed higher levels of flow because they had not only well-defined roles but also effective communication patterns and group dynamics. Previously, Bardzell, Bardzell, Pace, and Reed (2008) reported that for groups with “strong communication, tight social bonds, and competent role execution, player satisfaction rose” (p. 359). Finally, in this study, games teams found more success because gamers started the experience by selecting their own a unique role: the social networker, techie, pyro-technician, or photographer. Bardzell et al. (2008) previously found that when all team members were playing well-defined roles that they highly desired, teams garnered more success. .

Interdependency. In discussing teamwork in gaming, McGonigal (2011) explained that “we take great satisfaction in knowing we have a unique and important role to play in a much bigger effort” (p. 31). Thus, playing a role is not enough to support successful collaboration; the role needs to be unique and represent a piece of a larger effort. Previous research supports the effectiveness of mobile AR games to scaffold collaborative problem-solving through interdependent roles (Bressler & Bodzin, 2013; Dunleavy et al., 2009; Squire & Jan, 2007). Interdependent roles—such as the roles in *School Scene Investigators* as well as earlier INPLACE mobile games—were designed to be unique yet interconnected within the larger effort of the group.

Interdependency sets up the player to have a specific set of information to which he or she is the expert. In this study, gamers visited the crime scene and each role discovered something different. The social networker did not see any signs of forced entry and determined the thief must be at school. The techie noticed a white powder near the register and looked around for possible powder evidence in the cafeteria kitchen. The pyro-technician noticed all the money was gone, and the photographer documented the visual evidence at the crime scene by taking pictures. Players then shared their discoveries through oral communication. This concept of each player discovering different information continued throughout the game. Essentially, the game mechanic of interdependent roles set up the game teams to have more engaged communication patterns, and in turn, more positive social interactions.

Social interactions. There is actually a misconception that the general intelligence of a group's members predicts success; rather, researchers have shown that success is due in large part to the group's social interactions (Barron, 2003; Woolley, Chabris, Pentland, Hashmi, & Malone, 2010). In this chapter, I have already discussed how the findings from this study support and extend Barron's (2003) work on engaged and non-engaged responses and how communication responses connect to group success. This section will build on the work of collaborative problem-solving expert Anita Wooley and her colleagues. In 2008, Wooley, Gerbasi, Chabris, Kosslyn, and Hackman studied teamwork under four different collaborative situations: high expertise with and without collaborative scaffolding along with low expertise with and without collaborative scaffolding. They found that teams that had collaborative scaffolding were better at integrating information, regardless of expertise. In another study, Woolley et al. (2010) reported that team performance was partially predicted by an even distribution of conversational turn-taking. The next few sections will discuss how interdependent roles promoted effective social discourse and cooperation. Specifically, by incorporating the collaboratively scaffolding offered by interdependent game roles, the game helped to facilitate a more even distribution of conversational turn-taking and a more cooperative environment which led to greater success regardless of achievement level.

Discourse. In this study, as well as in previous research by Squire and Jan (2007), interdependency contributed to successful discourse. Squire and Jan (2007) found that the interdependent game roles of *Mad City Mystery* "encouraged students to share information, synthesize what they read, communicate orally with their group, ask questions, and debate meanings" (p. 24). Annemarie Palinscar (1998), an expert on social constructivist learning, explained that studies show students gain benefits from instructional discourse with peers: however, "the benefits depend upon the types of talk produced…it is important to attend to the structure of group activity so that responsibility is shared, expertise is distributed, and there is an ethos for building preceding ideas" (p.365). In this study, expertise was distributed via roles, responsibility was shared by distributing specific task and quests to each role, and the teams developed an ethos for building on ideas by using communal language and engaged responses. Game mechanics distributed responsibilities in such a way that facilitated the conversational turn-taking necessary for group success according to Woolley et al. (2010).

In this study, control teams did not benefit from interdependent roles; in fact, their roles were not defined for them at all and group dynamics suffered. In control teams, group difficulty seemed to stem from fighting over roles and responsibilities. Analytic performance suffered because the control teams without scaffolding—regardless of expertise—could not weigh each member's inputs appropriately (Wooley et al., 2008). Control teams did not evenly take turns speaking during their conversations; this was evident in the dynamics of all the control teams such as the ineffective democracy, egalitarian leadership, and dictatorship.

In game teams, interdependent roles set up successful communication patterns and conversational turn-taking and that led to more effective group dynamics. Group dynamics structures observed in the quasi-experimental group were community action, effective democracy, and expert guide. The team that represented above average math achievement was the community action team (Team E1); the four girls performed extremely well as a group in terms of their group discussions and decision-making. In the effective democracy, three boys and one girl were able to communicate effectively to get the job done but team members did not always support each other's ideas. Finally, in the expert guide scenario, the leader boy promoted a high level of discussion and instruction in this mixed gender group, and his ideas were rarely rejected by team members.

With the group dynamics and communication patterns so well established, game teams had more success at the performance-based transfer task. Depending on team pacing, game teams encountered a challenging performance-based task on Day 3 of Day 4. Teams needed to test the mystery powder. Every player had their role and the game dictated that the social networker should perform the vinegar test, the techie should perform the iodine test, the pyro-technician should perform the heat test, and the photographer should perform the pH test. During this experiment, game teams argued very little over roles and responsibilities. Instead, student teams concentrated on the task at hand, resolved confusion collectively, and got the job done efficiently. Continual communication certainly played a role in game teams' success at performing this task. Overall, successful conversational turn-taking derived from interdependency led to successful communication and collaboration, regardless of achievement level.

Cooperation. In this study, interdependency also contributed to a sense of cooperation. Interdependency is one of the elements of cooperative learning as defined by Johnson and Johnson (1999) along with individual accountability, face to face promotive interaction, social skills and group processing. As Klopfer (2008) pointed out in his book, *Augmented Learning*, the interdependent roles in mobile AR games are often designed using these components of cooperative learning as a design guideline. The game used in this study facilitated cooperation by incorporating some of the cooperative learning principles.

Positive interdependence. Johnson et al. (1993), who have extensively studied cooperation in the classroom, explained that "positive interdependence is successfully structured when group members perceive that they are linked with each other in a way that one cannot succeed unless everyone succeeds" (p. 9). In this study, gamers used more communal language revealing their perception that individual members were inextricably linked together as a group. In contrast, control teams used more command language showcasing their perceptions that each was an individual in the group, but there was not necessarily something linking them all together.

Individual accountability. According to Johnson and Johnson (1999), the next element of cooperative learning is that "students are held individually accountable to do their share of the work" (p. 71). In order to achieve this, SSI was designed using jigsaw pedagogy. The notion of jigsaw pedagogy is that each student in a group becomes an expert on one aspect of the activity, and teaches it to the other group members (Aronson, 1978; Aronson and Patnoe, 2011). As a team expert on a particular topic, each student was highly accountable to the group for his or her informational knowledge. In SSI, the social networker was an expert on the vinegar test; he or she had the knowledge of how each powder reacted to vinegar and what the reactions meant. In a similar regard, the techie was an expert on the iodine test, the pyro-technician knew about the heat test, and the photographer had all the information about the pH test. For the entirety of gameplay, each expert was accountable to the group for their expertise. In contrast, individuals on the control teams were not necessarily accountable to the teammates for any specific knowledge.

Promotive interaction. To accomplish face to face promotive interaction, students in groups need to do real work together and share resources, while supporting and encouraging each other (Johnson et al., 1993). First, in the game, students were doing the real work of investigators rather than the "busywork" of students. Second, as explained above, each player in the game had certain knowledge and information that they needed to share with the group; sharing was built right into their group dynamics structures, especially the community action team. The control activity was not designed to ensure this type of sharing. Lastly, since sharing was not a necessity, students in the control were not always encouraging each other, and in fact, outright discouraged each other at times with insults and command language. In contrast, game teams frequently used communal language—and accept and discuss responses—which demonstrated support and encouragement of teammate's ideas.

Group processing. Extending from promotive interaction, effective cooperative groups also engage in group processing. According to Johnson and Johnson (1999), group processing includes discussing goals for the group, making decisions together, and solving problems such that the group continues to function effectively. First, as exemplified in the case studies, game teams discussed their goals communally such as where they needed to go and what part of the report they needed to complete; only one of the control teams had similar discussions about goals. Second, case study teams from the game all made decisions together—it was a notable difference between game teams and control teams. Frequently, game teams made collective decisions about the next task to accomplish. Lastly, game teams resolved their confusion and worked through challenges collectively unlike the control teams. For the control teams, their inability to resolve confusion was one of the biggest problems in their group discussions. Game teams showcased a much stronger capacity for group processing than control teams and this was an important advantage.

According to cooperative learning experts Johnson and Johnson (1999), positive of interdependence largely dictates learning outcomes. Therefore, interdependency—as designed into the game in combination with other cooperative learning principles—led to cooperative social interactions which contributed to successful collaboration and deeper learning.

Learning together cooperatively. Johnson and Johnson (1999) contended that certain evidence will exist if educators have structured a cooperative learning activity properly. Students have to exhibit more effort to achieve, more positive relationships, and even shown signs of healthy psychological development. This type of evidence surfaced among game teams and supported that fact that cooperative learning was structured successfully in SSI.

Effort to achieve. Johnson and Johnson (1999) have found that students working well cooperatively are “motivated to continue to invest energy and effort in learning” (p. 72). This description bears striking resemblance to the ideas posited by Csikszentmihalyi (1990) about feelings of flow; people in flow are driven by intrinsic motivation to pursue their goals while continually upping the ante. As mentioned numerous times in this chapter, game teams experienced higher levels of flow than control teams which can be regarded as a sustained effort to achieve. Additionally, Johnson and Johnson (1999) stated that “active, involved students do not tend to engage in disruptive, off-task behavior” (p. 72). Since game teams experienced higher levels of flow—and had effort to achieve—they were more concentrated on the task and talked off-topic less than control teams.

Positive relationships. Since cooperative groups promote a sense of belonging, they also stimulate positive interpersonal relationships (Johnson & Johnson, 1999). One example of the burgeoning positive relationships among game teams was their frequency of engaged responses, offering more accept and discuss responses rather than reject responses. Interestingly, Jane McGonigal—a preeminent game designer—has made a similar comparison about gaming and positive relationships. According to McGonigal (2011), playing a game with someone can create a lasting social bond; in other words, playing games together promotes stronger social connectivity. In the first research study of the *School Scene Investigators* game model, we found that the collaborative scaffolding designed into the game seemed to help players build better relationships (Bressler & Bodzin, 2013). Therefore, the gameful learning and the cooperative learning offered by *SSI* model can both be linked to positive relationships between the players.

Psychological health and social competence. Since cooperative groups stimulate positive interpersonal relationships, they also have the potential to improve psychological health among the participants (Johnson & Johnson, 1999). This study did not capture explicit evidence of this phenomenon. However, one could argue that since game teams experienced higher levels of flow, they may have higher self-esteem because people in flow have clear, achievable goals since their skills are balanced with the challenge of the activity (Csikszentmihalyi, 1997). While empirical evidence of psychological health from this study may be slim, one teacher shared a story that can serve as anecdotal support. As explained in the case study of Team E3, the boy who took the lead in the Expert Guide team emerged as a truly confident and caring leader; he taught the rest of his group how to use the technology and helped them to understand the content and fill out their worksheets. This was a below average team and his assuredness was notable. In discussing this boy's game leadership with his teacher, there was a dichotomy between the boy's usual demeanor in class and during the game. According to the teacher, the boy is very quiet and shy in class. When he does speak, he often lacks confidence and stutters profusely. Yet, in the game, his confidence was apparent and his stuttering was minimal—possible evidence of improved psychological health and social competence that denotes an effective cooperative activity.

In summary, game teams were more successful at collaborating than control teams. I believe this result is best explained by the gamers playing interdependent roles: the game mechanic set up more successful communication patterns leading to a more enjoyable experience, more effective group dynamics and collaborative learning. Unfortunately, in control teams, undefined roles led to less successful communication patterns leading to a less enjoyable experience, and ultimately, less effective group dynamics and less collaborative learning. The next section will discuss the details of how the collaborative learning took shape and propose a theoretical model that brings together the findings from this study as well as previous research on learning and experiencing flow in groups.

Theoretical Proposition

In this chapter, I have discussed how and why game teams experienced higher levels of flow, more scientific practices in their reports and conversations, and more successful collaboration overall. In this section, I will synthesize all of this information into a theoretical proposition for how and why collaborative learning occurs in INPLACE mobile games.

As presented in Chapter 2, INPLACE is an acronym that stands for: **I**nterdependent, **N**etworked, **P**articipatory **L**earning, **A**ugmented, **C**ollaborative **E**xperience. INPLACE mobile games combine the best practices of what the games and learning field knows about how to engage students, how to support collaboration, and how to promote authentic practice in a discipline. Some notable characteristics of INPLACE mobile games include:

- a rich network of social interactions;
- participatory learning, specifically, situated learning;
- a cooperative framework via jigsaw pedagogy; and
- collaborative scaffolding and support via interdependent roles.

Reviewing these characteristics at this point is critical because they are prerequisites to promoting a successful group experience in INPLACE mobile games.

Earlier Work

My new model builds on the foundation of Johnson and Johnson's (1989, 1999) work in cooperative learning as well as the social flow research by Ryu and Parsons (2012).

Cooperation. According to Johnson and Johnson (1999), when a cooperative learning environment is structured effectively, individuals are accountable to the group through positive interdependency. Students working in such roles encourage and support each other (also called promotive interaction), as well as actively discuss goals of the group. These interactions result in not only learning benefits such as higher level reasoning, more idea generation, better learning transfer but also more concentrated time on task. Johnson and Johnson (1999) identified all the outcomes of cooperative learning and clustered them into three categories: effort to achieve, positive relationships, and psychological health/social competence. The entire interaction between the outcomes is laid out in Figure 17.

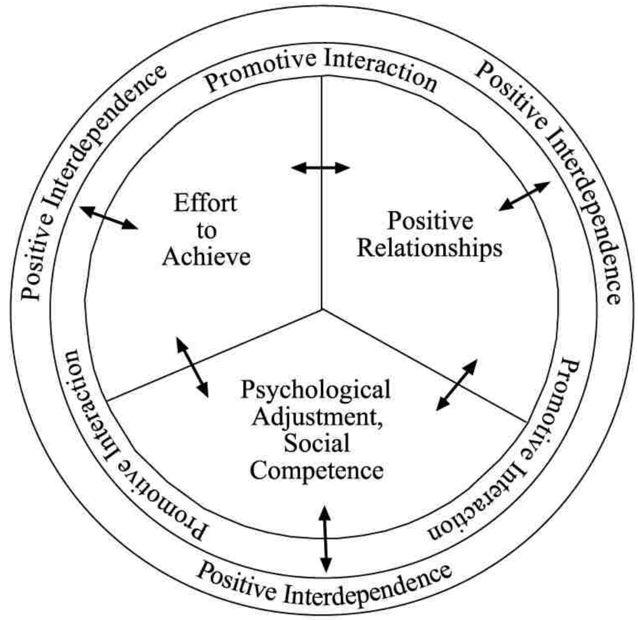


Figure 17. Outcomes of cooperation (Johnson et al., 1993).

My interpretation of Johnson and Johnson's (1999) findings is that—generally speaking—effective cooperation triggers social flow. When students are successfully cooperating, they experience increased concentration and clear goals; these are also two elements of flow. Furthermore, outcomes include improved effort to achieve and positive relationships; these are outcomes of an individual in flow. So a successful cooperative experience seems to trigger not only learning but also flow, or more specifically, social flow.

Social flow. Ryu and Parsons (2012) found that students participating in a collaborative mobile learning activity generated knowledge within the group; participants learned effectively by instantly sharing information generated by others. In their conclusion, they posited that the flow experience within the group, or social flow, promoted learning because those experiencing social flow challenged each other. Succinctly, social flow acts as a trigger for the learning (Ryu & Parsons, 2012). The authors created the representation shown as Figure 18 to depict that social flow pushes people to tackle challenges and this results in collaborative learning.

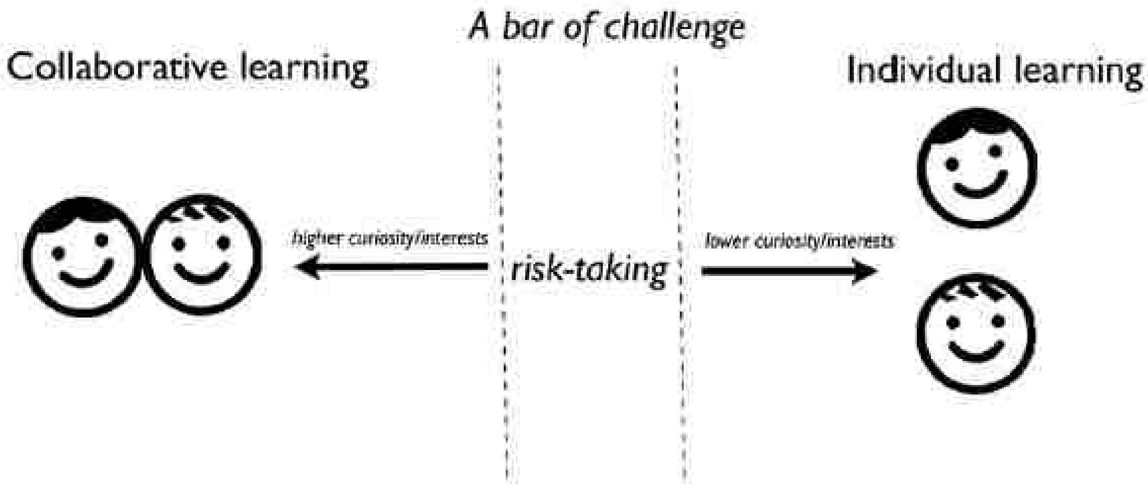


Figure 18. Social flow that raises the bar of challenge (Ryu & Parsons, 2012).

Additionally, according to Ryu and Parsons (2012), “social flow may make explicit why learners would cooperate and how it can be strengthened by collaboration” (p. 712). In other words, social flow may not only trigger collaborative learning but also the experience of such an effective collaboration may reinforce a group’s social flow. According to Walker (2010), who posited that social flow is a different phenomenon from individual flow, explained that “in highly interdependent situations, people may serve as agents of flow for each other” (p. 4).

Model Introduction and Explanation

In previous mobile AR games, the use of interdependent roles has shown evidence of collaborative learning. Specifically, this study showed that the use of interdependent roles in the quasi-experimental intervention—an INPLACE mobile game—scaffolded the social interaction and discourse necessary to not only build scientific knowledge but also promote an intrinsically rewarding experience. Presented in Figure 19 is my theoretical proposition for how cooperation, social flow, and collaborative learning work together to promote a successful group experience in INPLACE mobile games.

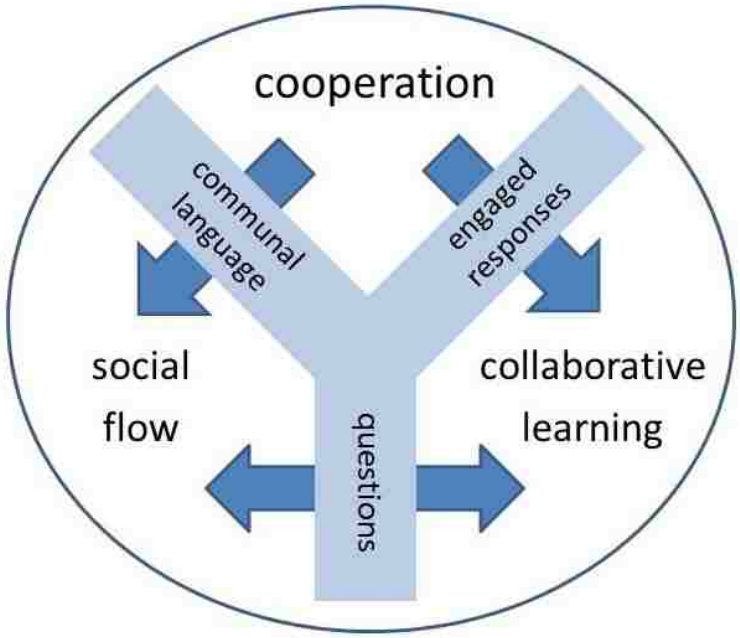


Figure 19. Social contexts and communication within INPLACE mobile games.

Through this chapter, a major theme has been that the primary driver of the learning experience offered by INPLACE mobile games is the conversational discourse and communication between group members. In brief, game teams learned by talking to each other effectively. Therefore, the model—as shown in Figure 19—is held together by several communication channels each identifying a prominent discourse style that is found there. Essentially, communication is the glue that keeps the team together and functioning successfully.

The three social contexts that group members shift between are cooperation, social flow, and collaborative learning. These contexts are not mutually exclusive; they are interconnected facets of a successful group experience. The arrows indicate triggers between these social contexts. Previous research has identified cooperation as a trigger for both social flow and collaborative learning, while social flow and collaborative learning reinforce each other. Each context is defined by a set of interactions and feelings. The following subsections describe how communication unifies the experience that continues to shift between these three basic contexts.

Defining cooperation, collaboration, and social flow. Cooperation and collaborative learning are often used interchangeably; however, the terms are separated in this model. As defined by Dillenbourg (1999), "in cooperation, partners split the work [and] solve sub-tasks individually…In collaboration, partners do the work 'together'" (p. 8). Dillenbourg's succinct explanation of the division of labor is precisely the reason that the terms are applied to different social contexts in the model. In the model, *cooperation* represents students doing their individual work as a piece of the group's mission. This includes feeling of positive interdependency and individual accountability generally achieved through jigsaw pedagogy. In *collaborative learning*, students are working together to construct new knowledge. They might be collectively discussing new ideas, trying to resolve confusion, or working through intellectual challenges. Finally, in *social flow*, the group is experiencing flow together. According to Walker (2010), to achieve social flow the group needs (a) a team-level challenge that is significant, such as solving the mystery of who stole the money from the cafeteria and left behind a mysterious powder, (b) to agree on their goals and roles—this capability was built into the game through quests and game roles, and (c) patterns of harmonious interpersonal relations as was showcased through the game teams' group dynamics.

Cooperation and social flow. The communication pattern that seems to unify these two social contexts is communal language. When an individual feels like a critical member of the group, he or she uses communal language to help guide the group towards a goal. The use of communal language then promotes harmonious interpersonal relations crucial for social flow.

Cooperation and collaborative learning. The communication pattern that seems to unify these two social contexts is the use of engaged responses, such as accept and discuss responses. When an individual is working on a sub-tsk for the group, he or she uses discuss responses to share discovered information with the group. Group members, in turn, use accept and discuss responses to work together on constructing new knowledge.

Social flow and collaborative learning. The communication pattern that seems to unify these two social contexts is the use of questions. In this study, game teams asked a lot of questions of each other. First, some questions sought clarification of an individual's sub-task. These questions seemed aimed to gather information from other individuals so that the group could learn from each other. Second, a subset of questions sought to confirm or establish the group's current goal, so that all members of the group were thinking in a similar way. Lastly, a different subset of questions sought to resolve confusion or challenge the group to construct new knowledge. When groups were challenged or confused, they managed the disequilibrium effectively through group dialogue. Communication patterns and dynamics enabled them to overcome such frustration and re-establish social flow.

Where is teacher involvement? This is a model to explain how students are interacting with each other and learning from each other within an INPLACE mobile game. As discussed in Chapter 4, game teams required very little teacher facilitation of their learning; game teams were autonomous, self-directed, and on task. As summarized by Kirriemuir and McFarlane (2004), "games are often a facilitator to social, communication and peer activities…learners gain resources from fellow members that equip them to solve problems within" (p. 15-16). Therefore, teacher involvement was specifically left out of the model. When game teams did include the teacher in a discussion, it was usually through a line of questioning in which the teacher functioned as a collaborative member of the group.

CHAPTER 6: IMPLICATIONS AND RECOMMENDATIONS

This study began with the premise that INPLACE mobile games have the potential to promote flow experience, scientific practices, and successful collaboration among teams of players. I found game teams experienced not only higher levels of flow but also higher levels of scientific practices than control teams. My findings revealed that game teams collaborated more successfully than control teams because gamers communicated with engaged responses and communal language; control teams communicated with higher levels of rejecting responses and command language which generated ineffective group dynamics. I concluded that the game design was largely responsible for the difference: playing interdependent roles promoted a cooperative ethos among game teams. The game also revealed additional learning benefits such as reduced off-topic conversations, reduced confusion, and fewer classroom management issues which enriched the role of the teacher by enabling them to offer more directed assistance to small groups.

This study culminates over 27 months of design, development, pilot field-testing, revisions, a scaled-up implementation, data collection and analysis and—ultimately—the complete results and conclusions. I began this study aiming to find a way to deeply engage students in science learning that would be practical given the parameters of a school setting. I have come away with not only a scalable model that demonstrated such an achievement but also a more profound understanding of the relationships between the game mechanics and how they promote social flow and collaborative learning. Given the expanded insight that this study has offered, I start this chapter with the overall significance of the study and then follow with discipline specific implications. Then, I present the limitations of the study design and the findings. In the next section, I offer my case for INPLACE mobile games as a practical application of game-based learning in formal education that satisfies several barriers to implementation. In the penultimate section, I discuss two major directions for future research. Lastly, I conclude with a reflection on the relationship between formal education and INPLACE mobile games.

Significance of Study

This dissertation study supports the notion that doing things *differently* can mean doing things *better*; the key is to disturb the inertia of the classroom. Recently, Norris and Soloway (2011) recently reminded the educational community that "as long as schools use computers to teach the existing curriculum using existing pedagogies little to nothing will be gained" (p. 4). Back in the 1990s, the Apple Classrooms of Tomorrow (ACOT) demonstrated not only marked increases in student scores but also enhanced skills in inquiry, collaboration, and problem solving…well beyond traditional classrooms (Dwyer, 1994). The ACOT showed such differences because they let the technology ‘disturb the inertia’ of the traditional classroom; they did not teach existing curricula with existing pedagogies. Instead, ACOT let the affordances of the computer drive pedagogical and curricular changes—and they garnered results.

Similar to ACOT, this INPLACE mobile game disturbed the inertia of the classroom. INPLACE mobile games are designed around the affordance of mobile technologies and the “gamefulness” (Betts, 2013) of the educational content; they do not deliver existing curricula. By disturbing classroom inertia in these ways, INPLACE mobile games are a step in the right directions toward changing the culture of school. As Squire (2005) indicated, “the real challenge is not so much in bringing games—or any technology—into our schools but rather changing the cultures of our schools to be organized around learning instead of the current form of social control” (p. 5). As this study demonstrated, game teams developed and exhibited a *deeper* understanding of scientific practices and needed *less* teacher intervention to stay on task. In other words, the game teams in this study required less teacher-centric social control to achieve better learning results. Old models of learning where the teacher delivered instruction, maintained control, and the student acquired content worked well enough when the aim of education was content mastery. However, as Richardson argued (2012), our end goal as an educational system can no longer be content mastery—it needs to be replaced by learning mastery. INPLACE mobile games can help reform education from content mastery to learning mastery since they exemplify the case and point that relinquishing control to the learner actually improved their learning.

Discipline Specific Implications

Since this study was structured around the three themes of scientific practices, flow, and collaboration, in this section I will identify the implications for each of those disciplines as well as game-based learning in general.

Science Education

Previous research set the bar and determined that collaborative science games played on mobile devices held promise to improve science learning. This study provided further evidence to distinguish the science learning advantages offered by an INPLACE mobile game over a traditional approach. One of the main reasons that meaningful learning took place during the game was due to the situated, immersive context. By role-playing within the game, students were learning and living scientific practices. Specifically, students were not learning science by simply acquiring facts. Rather they were immersed in a context through which science became a process, and more importantly, the students were eager to participate in that process.

Unfortunately, while the game was designed to offer a more authentic, situated learning experience, it was not *perceived* as a real science experience by all the students. After completing the game, one student said: "This was actually really fun. I don't want to go back to doing real science." According to the report commissioned by the NRC (2012a), today's science students have developed the misconception that science is about memorizing isolated facts because there has been a rigid focus on solely content for too long. The boy in this study presumably was working from such an assumption that memorizing facts is real science; therefore, the game could not be real science.

Clearly, science education needs a "marketing make-over." As shown in this study, the situated role-playing of an INPLACE mobile game affords students opportunities to experience science as a fun, cohesive exploration and not simply as facts to memorize. If *Civilization III* enabled students to think of history as more than just isolated facts (Squire, 2004) than INPLACE mobile games can help students view science as more than content memorization. We may be able to adjust the impressions that young people have about science if we design and implement more INPLACE mobile games.

Maybe this adjustment would also convince more students to enter STEM disciplines. When students actually do real science, it can "pique students' curiosity, capture their interest, and motivate their continued study" (NRC, 2012a, p. 42). Once we get the students interested in science through such games, continuing to offer similar interactive experiences may also keep them retained in STEM disciplines as they enter higher education; according to recent article published in the *Journal of College Science Teaching*, researchers claimed that class activities that promote thinking, responding, and interacting can retain those in STEM majors (Watkins & Mazur, 2013).

Computer-Supported Collaborative Learning

To truly satisfy a science marketing make-over, students need to perceive science as a 'social enterprise' and they need to learn to collaboratively solve problems for success in scientific careers (NRC, 2012a). Unfortunately, students are not prepared for such collaborative work because school environments place too much emphasis on individual work (Bransford et al., 2000) and standardized testing (Ravitch, 2011).

As discussed in Chapter 2, well-designed multi-player games foster effective collaborative learning and social construction of knowledge. In this study, there was evidence of successful collaboration among game teams as documented by the higher levels of scientific practices, engaged responses, and communal language present in their conversations. Even though the collaboration was deemed successful, individuals need to be held accountable for their own learning to fit within today's school structure. My study put students in a collaborative problem solving environment yet evaluated each player with an individual assessment.

INPLACE mobile games—one of the newest iterations of computer-supported collaborative learning—offer a unique way for students to socially construct knowledge, while also attaining their own individual learning goals. INPLACE mobile games, such as the one used in this study, fit the image of science as a social enterprise as described in the NRC (2012a) report. Incorporating more INPLACE mobile games into classrooms could better prepare students for the collaborative culture of their future careers while also maintaining individual scores to align to the formal grading procedures of most schools.

Flow, Games, and Learning

While individualized assessment procedures and improved scientific practices satisfy administrators and parents, the actual gameplay that takes place within an INPLACE mobile game satisfies the students. As mentioned in the first few chapters, improving engagement is important for reducing drop-out rates (NRC, 2011b). Students on game teams had higher levels of flow than control teams. Since flow is related to engagement, game teams were more engaged. INPLACE mobile games may be a way to re-engage students in school…and learning.

Specifically, the traditionally low-performing students may become more engaged in school through INPLACE mobile games. Research shows that game-based curriculums are "successful, if not more so, at supporting science learning among students classified in lower academic levels than their peers participating in higher academic level courses" (Sadler, Romine, Stuart, & Merle-Johnson, 2013, p. 493). This may be due to the deeper engagement that low-performing students experience during game-based learning. According to the report commissioned by the NRC (2011b), students with low-achievement—such as those represented by Team C3 and Team E3 in this study—may be low-performers as a result of years of disengagement with school. This study showed that, beyond higher levels of flow, the low-achievement students on the game team (Team E3) had notable differences from the low-achievement students in the control (Team C3). First, Team E3 showcased levels of scientific argumentation on par with the higher achievement game teams. In contrast, Team C3 could not showcase this practice without teacher intervention. Second, Team E3 went off-topic only a few times over the whole activity, whereas off-topic discussions occurred within Team C3 frequently and for long periods of time. Lastly, the language style of Team E3 was highly communal with very low rejecting responses; therefore, they collaborated effectively. Unfortunately, the language style of Team C3 was rarely communal and frequently included rejecting responses; therefore, their collaboration was not as effective. Overall, it seemed that the game environment was not only more engaging for low-achievement students but also more effective at keeping them on-task and collaborating within the learning experience.

This study also shed light on the relationship between flow and learning for all levels of students. Previously, the flow state has been connected to positive affect and improved athletic performance; this study showed that it is also connected to higher learning levels. Higher levels of flow during gameplay were associated with higher learning levels, for all levels of students in both the experiment and control groups. We still need future studies to further understand this relationship.

This study indicated that social interactions may be a complicating factor, particularly social flow. Quantitatively, this study concluded that game teams had higher levels of flow and generated higher report scores. Qualitatively, the game teams seemed to be better collaborators because of their social flow. Thus, social flow is connected to the development of more scientific practices and successful collaboration. As mentioned earlier in this dissertation, to make the U.S. globally competitive in science and technology, students need to be engaged with science education, engage in authentic scientific practices, and learn to collaborate successfully; findings from this study show that an INPLACE mobile game can meet all of those requirements. Since this game model promotes flow and improved science learning, perhaps it would help improve engagement and learning in other subject areas as well.

Game-Based Learning

This study has added to the body of evidence that supports game-based learning; the INPLACE mobile game used in this study supported higher scientific practices in both written reports and collaborative conversations. According to a report commissioned by the NRC (2011a), games show potential to be effective for science learning, but the results to date are inconclusive at worst and limited at best. To create a stronger knowledge base, researchers need to specify learning outcomes and discuss the underlying theory as well as the design features that activate learning (NRC, 2011a). This dissertation has addressed each of the components at length: the learning outcomes of the game were scientific practices, the underlying theories were social constructivism and situated cognition, and the design features that activated the learning were interdependent roles. Hopefully future research studies will follow suit and construct studies that integrate these components in order to further strengthen the knowledge base.

For the games and learning community, it is also worth mentioning the significance of the *mobile* in the category of INPLACE mobile games. Recently, Wouters et al. (2013) published a meta-analysis of serious educational games that revealed they were no more motivating at improving learning than passive instruction. The findings from this study are contradictory because game teams had higher flow, which includes motivation to learn, and they had higher report scores. Could *mobile design* make the difference?

Wouters et al. (2013) did not investigate the variable of mobility; however, they did determine that participants who played serious educational games in a group learned more than participants who played alone. Their conclusion provides additional support for the promise of INPLACE games since the design is inherently collaborative—it's a prerequisite design element in order to be considered INPLACE. However, INPLACE mobile games also take advantage of the affordances of mobile devices. Specifically, mobility enables learning to become more contextualized and situated. The games offer an immersive experience within a real, authentic context which enhances learning transfer. The physicality of working with mobile devices also increases players' enjoyment. Overall, the game and learning community should continue to take a closer look at gaming with mobile technology. The wide-spread use of mobile technologies in school (Johnson et al., 2013) and their inherent affordances may make collaborative mobile games a viable way to save American schools from obsolescence.

Limitations of Study

There were several limitations to this design of this study. First, the post-test-only control group design was a limitation. Without a pre-test to justify that the groups are equivalent prior to the intervention, classes were randomly assigned to control and experiment groups to achieve probabilistic equivalence and math PSSA scores were used as a covariate to make group comparison fairer. Second, the sample was obtained from only one urban middle school; therefore, the generalizability of the findings is somewhat limited. Care must be taken when generalizing to other contexts. Third, the control activity was taught by two different teachers. While they were very similar in their teaching style, they had difference classroom management abilities. One teacher was a veteran teacher with over 10 years in the classroom and clearly demonstrated classroom management strategies; the other teacher was in his second year of teaching and his classroom management skills were not as strong as the seasoned teacher. To minimize instructor differences, each teacher taught some control and some experiment classes. Fourth, the intervention was very short is duration. A few days in the course of a student's life may not significantly impact learning outcomes.

Another significant study limitation was the written learning outcome. First, a piece of writing is inherently impacted by the student's writing ability. This type of written analysis may have been difficult for some students. Such students may have garnered the benefits of the activity but may not have been able to write about them due to a language barrier or some other writing impairment. To make group comparison fairer, reading PSSA scores were used as a covariate in the second MANCOVA analysis. The case study research was also included in order to highlight learning that took place in discussions but that may not have been reflected in written reports.

Additionally, while the MANOVA results did produce statistical significance, it did not garner a large effect size. As noted above, the written learning outcome may not have captured the truest measure of learning for those with a lower level writing capability; however, it also may not have captured the complete learning outcome of gamers. Students playing the INPLACE mobile game had to fill out a paper-based worksheet separate from the game. This meant they had to break the 'magic circle' (Salen & Zimmerman, 2004); they stopped being *investigators solving a crime* and switched to being *students filling out a packet*. This clearly frustrated some players as showcased by this comment: "I hate the darn packet" (*Team E1 Conversation, 532*). Since the capacity for situated learning is one of the game's affordances that generated meaningful learning, leaving that situation may have strained the learning outcome. Moreover, leaving the magic circle multiple times may have interrupted the gamers' flow experience generated in part by the immersive feel of the game. If the flow scores and learning scores of gamers had been better representations, maybe the effect size would have improved.

Lastly, the participating teachers and the school's principal were extremely supportive of this research project. This meant that the researchers were welcomed into the building and were able to freely disturb the inertia of the classroom as well as the hallways of the school—a necessity since the students on the game teams were in the hallways for much of the activity. Students on the game teams may not have achieved markedly higher levels of engagement and learning outcomes working with less enthusiastic teachers or within a school that had less supportive administration.

Practical Application

Even with such limitations, the study still offers a practical application for formal education. As I will discuss in this section, the implementation costs are low and the game fits the logistics of a school setting while offering an avenue to enrich the teacher's role. The game designed for this study also meets and exceeds curriculum learning objectives for today's science education. Most importantly, educators can use the research results from this study to support their case to administrators for why they want to use INPLACE mobile games for science teaching and learning.

To begin, INPLACE mobile games are an appropriate consideration for educators given the significant challenges K-12 education is facing according to the *2013 Horizon Report*. Johnson et al. (2013) explained that students are still learning in a lecture-then-test setting yet non-traditional alternatives are needed; specifically, "in order for students to get a well-rounded education with real world experience, they must also engage in more informal in-class activities" (p. 10). Through situated learning and collaborative gameplay, INPLACE mobile games offer an effective, non-traditional alternative to this challenge.

Going Beyond Current Curriculum into the Next Generation

Even when teachers have enthusiasm for game-based learning, there are always concerns from parents and administrators that educational games do not map well to curriculum objectives (Kirriemuir & McFarlane, 2004). Not only did this study map directly to the curriculum objectives, it was adopted by the district as an accepted curriculum unit and statistically proved to offer a more engaging learning experience with higher learning payoffs than the traditional, control activity.

This game and the accompanying worksheet also fit the current rhetoric of accountability; student learning was assessed using a paper and pencil worksheet that was graded with a rubric aligning to Next Generation Science Standards (NGSS). With the written incident report, teachers can concretely demonstrate the outcome of student learning to parents.

Game-based learning has struggled to enter classrooms because parents express concerns over using "games in lesson time since such skill development [does] not match the criteria assessed in high stakes national tests" (Kirriemuir & McFarlane, 2004, p. 19). To further alleviate parental concerns, this INPLACE mobile game exceeds beyond basic curriculum standards and effectively meets the needs of NGSS. With the recent release and adoption of the NGSS, schools will need curriculum activities that support student learning aligned to these standards. The culminating report was perfectly aligned to the NGSS and the study showed that the game players had higher report scores.

Beyond the scientific practices, educators can also refer to the 21st century learning skills of communication and collaborative problem-solving as highly desirable outcomes of gameplay. These skills are not cultivated in the majority of U.S. schools, yet they are essential to garnering success as adults (NRC, 2012b). In this study, game teams had better communication and positive social interactions; therefore, they were able to share decision making and co-construct solutions. It is particularly unique in that students worked *collaboratively* but were assessed *individually*.

Reduced Barriers to Implementation

Overall, educators have an array of curricular reasons to justify use of this INPLACE mobile game for parents and administrators. This study demonstrated that the game not only afforded higher levels of scientific practices but also aligned to NGSS and individually assessed learning within a collaborative activity. Another reason that INPLACE mobile games should be considered by school policy makers is the reduced barriers to implementation. This study addressed implementation issues as mentioned in Chapter 2 including implementation costs, the constraints of the school day, and the effort required of teachers to keep kids on track.

Implementation costs. As discussed in Chapter 2, the first barrier to implementation was implementation costs such as the costs of licenses, software products, and video game equipment (Kirriemuir & McFarlane, 2004). In order to ensure practical application, this study used a freely available iPad application and took advantage of the iPad cart and wireless internet that was already available at the school. There were no implementation costs carried by the school or the researcher. While some might see an iPad cart as a sizable school investment, the reality is that many schools are using iPads; as of June 2013—according to Apple—there were over 10 million iPads in schools and the number is steadily growing (Leonard, 2013). Additionally, more and more schools are inviting students to *bring your own device (BYOD)* which include various mobile devices from tablets to phones (Johnson et al., 2013). If schools do not have an iPad cart, they could consider a BYOD program to keep implementation costs low.

At one time, educators opposed the idea of students using mobile devices in school; however, according to the *2013 Horizon Report*, that opinion has shifted. Johnson et al. (2013) explained that we are on the verge of widespread adoption in large part due to tablet usage, such as iPads. INPLACE mobile games offer a specific learning activity for iPads, but these games will also benefit from the widespread adoption of tablet use in schools. In a 1:1 iPad school, where students all have their own tablets, time would not be wasted at the beginning of the activity to hand out the devices and explain how to use them. It might also decrease the amount of time spent addressing technical issues because the students would be more accustomed to working with the devices.

Role of the teacher. The second barrier to implementation was that teachers struggle to keep kids on track and gaining the appropriate educational experience during gameplay (Kirriemuir & McFarlane, 2004). In the study, off-task behavior was more prevalent among control teams than game teams. Teachers had to work harder to manage the classroom behavior of control teams. In contrast, game teams required less teacher intervention, yet remained on-task, encountered less confusion, collaborated more successfully, and achieved higher levels of scientific practice in both their conversations and reports.

With the INPLACE mobile game, the role of teacher naturally shifted from *sage on the stage* to *guide on the side*. Curriculum activities such an INPLACE mobile games allow the teacher to provide targeted assistance to students as they need help in attaining learning mastery. This role change fits with one of the key trends happening in K-12 education according to the *2013 Horizon Report*. Johnson et al. (2013) explained that educators need to revise their roles in order to best prepare students for the world in which they will live and work; specifically, the abundance of easily accessible and feely available information is “challenging us to revisit our roles as educators” (p. 9).

Constraints of school day. The third barrier to implementation discussed in Chapter 2 was that students could not gain maximum benefit from serious educational games due to the constraints of the school day (Young et al., 2012). In order to ensure practical application, this study took place during regular class meeting times using school resources within the building. As demonstrated by this study, INPLACE mobile games can be broken down into chapters and align to the standards and completed across several class periods.

Directions for Future Research

In this section, I would like to focus on two major areas for future research. First, given the NGSS, I believe some additional research could be conducted to refine the game to better assess learning and promote scientific discussions. Second, given the description of science as a social enterprise from the NRC (2012a) report, and given the practical application of this study, future research could investigate how students' learning progresses from thinking like students to thinking like scientists.

Game Redesign

Game mechanics were partially responsible for individual and social flow, as well as promoting scientific practices. Future research on the game design could improve the game's ability to afford flow and learning.

Flow. Since flow has a small predictive relationship with learning it becomes relevant to better understand the game mechanics that maintain and support player's flow as well as those that diminish a player's flow.

This study has shown that a well-designed mobile science game promotes higher levels of flow and higher levels of learning as demonstrated on the culminating artifact measure. Not only are INPLACE mobile games engaging, but when they are well-designed, the engagement can be quantified according flow theory. Future research could study the game design principles that are most effective at promoting flow.

Future research should also seek to alleviate problems where players leave the flow state. As mentioned earlier, game teams had to break out of the magic circle (Salen & Zimmerman, 2004); leaving the magic circle undoubtedly has a negative impact on a player's flow experience. The biggest reason for leaving the magic circle was filling out the worksheet. Game redesign should include a more seamless assessment that keeps the magic circle intact. The ARIS programming environment does provide a virtual notebook which was not utilized in this game version. Perhaps this notebook feature—since it is within the gaming environment—could serve as a better assessment tool and be built into the next version. The flow state of students playing the game using the old assessment versus those using the new assessment could be compared to determine if the new assessment affords a higher level of flow.

Scientific practices. Game mechanics were partially responsible for promoting scientific practices. Re-designing aspects of the game's narrative could help to improve each player's ability to plan the investigation and argue with evidence. Such redesigns could also improve players' flow experiences by immersing them more deeply in the narrative and protecting the experience of the magic circle.

Planning the investigation (practice #3). In reviewing student dialogue, control teams had a better understanding of the plan they needed to execute in order to determine the identity of the mystery powder. In order to improve the ability of game teams to plan out their investigation, some game redesign is needed. Currently, students were apt to plan their investigation in terms of the game narrative. For example, student might state that they needed to interview suspects rather than test the known powders.

The narrative is clearly a powerful way for players to comprehend their goals; unfortunately, they needed to plan out how to find out the mystery powder's identity which was a bit buried in the narrative. Therefore, game redesign should re-evaluate the narrative and make the mystery powder and the known powders a more prominent part of the narrative. In order to determine narrative revisions, mocking up prototypes and then conducting focus group research with smaller groups of students to better understand their thought processes as they progress through the game and try to plan out their investigation would make sense.

Arguing with evidence (practice #7). In reviewing team dialogue, game teams all revealed moderate levels of Practice #7 while control teams had very low levels. While the game was clearly better at promoting this practice, it could still be redesigned to better facilitate this practice. For Team E3, the below average game team, their evidence-based arguments revolved entirely around the story, rather than the powder evidence. Re-evaluating the narrative and integrating more information about the powders might help to alleviate this problem. Yet, this practice is probably one of the most difficult scientific practices for the teams to achieve. Perhaps this game primes the students to start constructing explanations and thinking through the evidence. Yet, to extend the learning of this practice, another similarly designed game needs to take the students deeper into this practice as they work to solve another evidence-based mystery.

Learning Progression from Student to Scientist

Based on the findings of the pilot study (Bressler & Bodzin, 2013) and the findings from this dissertation game, future research could use a series of INPLACE mobile games and investigate how students' learning progresses from thinking like students to thinking like scientists. According to a report published by the NRC (2007b), over years of science education, students need to successively develop more sophisticated ways of thinking about science concepts and practices; this process is called a learning progression. Several games in a series could build on each other and scaffold a specific learning progression. As the students move from game to game, the challenges would get harder, in order to keep them in flow, and less scaffolding would be provided for their discussions in order for players to construct deeper knowledge on their own using their own ideas.

Since there is already an existing body of INPLACE mobile games—such as *Environmental Detectives*, *Mad City Mystery*, *Outbreak @ The Institute*, and *Alien Contact*—perhaps an educational grant could be written to unite all these research projects under one science learning progression. All of these games integrate interdependent roles. One week a student might work with her group as a chemist trying to determine why aliens landed in the school yard, whereas the next week she can become a medical technician trying to determine the cause of a disease outbreak. In each scenario, students would get exposed to different science concepts, careers, and ultimately build a science learning progression over time.

As a derivative work of learning progressions, it will also be important to understand how well these games work for all achievement levels and, if they work well for low performing students, why is that the case. In this dissertation study, the team with the most individual educational plans (IEPs) collaborated and communicated quite successfully while also staying focused on the task. INPLACE mobile games might be of interest to the educational psychologists who study attention-deficit disorder (ADD), attention-deficit/hyperactivity disorder (ADHD), and other attention disorders. There is reason to believe that INPLACE mobile games could have a positive impact on students with ADD and ADHD. Students with ADHD showed improved performance in gaming over a basic feedback scenario because gaming helps overcome the deficits caused by the executive functions of the ADHD brain (Dovis, Oord, Wiers, & Prins, 2012).

Last words

The 2011 report published by the NRC advocated for further investment in game-based learning, specifically as an avenue to improve science learning. The report asserted that games have potential to achieve many goals of science education including cultivating motivation to learn science as well as scientific argumentation and discourse. This study confirmed the potential of INPLACE mobile games to cultivate motivation by promoting higher levels of flow and deeper levels of scientific practice than traditional science education.

This study designed and implemented a new style of curriculum activity that does not represent business-as-usual instruction; this INPLACE mobile game disturbed the inertia of the classroom and demonstrated not only improved learning but also high engagement. We need to create and enact curriculum activities that ensure students will garner important skills and benefit from the unique affordances of mobile technology—no more business-as-usual. As Richardson (2012) explained, the new paradigm of education should be about "asking questions, working with others to find the answers, [and] doing real work" (loc. 267). That is exactly what students are doing when they are playing an INPLACE mobile game.

INPLACE mobile games can satisfy this new educational paradigm and—what is most intriguing—they can satisfy the concerns of both adults and students. Administrators want low costs; INPLACE mobile games are essentially free. Teachers want easy implementation with no added work; the game used in this study required minimal professional development and was easily blended with an existing activity. Parents want concrete evidence of standards-based learning; individualized assessments can be designed to provide proof of game-based learning. Students want active, exploratory learning; INPLACE mobile games not only offer this style of learning but also challenge, engage, and—for all intents and purposes—empower students.

Are INPLACE mobile games the perfect panacea to solve all of the issues currently facing our educational system? No, because nothing can do that. However, there is a strong case that they can improve science learning and engagement which means they can address two profound issues facing education today. Ultimately, by serving the needs of administrators, policy makers, educators, and parents along with the desires of the students, INPLACE mobile games are a win-win. In other words, they are the perfect “win state” for educational reform. ☺

References

Achieve, Inc. (2013). *Next Generation Science Standards*. Washington, DC: Achieve, Inc. Retrieved from http://www.nextgenscience.org/

Admiraal, W., Huizenga, J., Akkerman, S., & Ten Dam, G. (2011). The concept of flow in collaborative game-based learning. *Computers in Human Behavior, 27*(3), 1185-1194.

Anderson, M. A. (2012). What is it about The Oregon Trail? Letting learners 'drive'. *EventDV, 2012*(5), 20-22.

Annetta, L. A., & Cheng, M. (2008). Why educational video games? In L. A. Annetta (Ed.), *Serious educational games: From theory to practice* (pp. 1-11). Rotterdam, The Netherlands: Sense Publishers.

Annetta, L. A., Minogue, J., Holmes, S. Y., & Cheng, M.-T. (2009). Investigating the impact of video games on high school students' engagement and learning about genetics. *Computers & Education, 53*(1), 74-85.

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.

Arafeh, S., Levin, D., Rainie, L., & Lenhart, A. (2002). *The digital disconnect: The widening gap between Internet-savvy students and their schools*. Washington, DC: Pew Internet & American Life Project. Retrieved from http://www.pewinternet.org/~/media//Files/Reports/2002/PIP_Schools_Internet_Report.pdf.pdf

ARIS. (2013). *Mobile learning experiences – Creating educational games on the iPhone*. Retrieved May 16, 2013, from http://arisgames.org/

Arnab, S., Berta, R., Earp, J., de Freitas, S., Popescu, M., Romero, M., . . . Usart, M. (2012). Framing the adoption of Serious Games in formal education. *Electronic Journal of e-Learning, 10*(2), 159-171.

Aronson, E. (1978). *The jigsaw classroom*. Beverly Hills, CA: Sage Publications.

Aronson, E., & Patnoe, S. (2011). *Cooperation in the classroom: The jigsaw method* (3rd ed.). London: Pinter & Martin Ltd.

Asbell-Clarke, J., Edwards, T., Larsen, J., Rowe, E., Sylvan, E., & Hewitt, J. (2012). Martian Boneyards: Scientific inquiry in an MMO game. *International Journal of Game-Based Learning*, 2(1), 52-76.

ATLAS.ti. (2013). *ATLAS.ti: Qualitative data analysis* (Version 6.2.28) [Computer software]. Berlin, Germany: Scientific Software Development GmbH.

Banathy, B. H. (1994). Designing educational systems: Creating our future in a changing world. In C. M. Reigeluth & R. J. Garfinkle (Eds.), *Systemic change in education* (pp. 27-34). Englewood Cliffs, NJ: Educational Technology Publications, Inc.

Barab, S. (2006). Design-based research: A methodological toolkit for the learning sciences. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 153-169). New York: Cambridge University Press.

Barab, S., & Dede, C. (2007). Games and immersive participatory simulations for science education: An emerging type of curricula. *Journal of Science Education & Technology, 16*(1), 1-3.

Barab, S. A., Sadler, T. D., Heiselt, C., Hickey, D., & Zuiker, S. (2010). Erratum to: Relating narrative, inquiry, and inscriptions: Supporting consequential play. *Journal of Science Education & Technology, 19*(4), 387-407.

Barab, S., Scott, B., Siyahhan, S., Goldstone, R., Ingram-Goble, A., Zuiker, S., & Warren, S. (2009). Transformational play as a curricular scaffold: Using videogames to support science education. *Journal of Science Education & Technology, 18*(4), 305-320.

Barab, S., Zuiker, S., Warren, S., Hickey, D., Ingram-Goble, A., Kwon, E.-J., . . . Herring, S. C. (2007). Situationally embodied curriculum: Relating formalisms and contexts. *Science Education, 91*(5), 750-782.

Bardzell, S., Bardzell, J., Pace, T., & Reed, K. (2008). *Blissfully productive: Grouping and cooperation in World of Warcraft instance runs*. Paper presented at the Proceedings of the 2008 ACM conference on Computer supported cooperative work, San Diego, CA, USA.

Barendregt, W., & Bekker, T. M. (2011). The influence of the level of free-choice learning activities on the use of an educational computer game. *Computers & Education, 56*(1), 80-90.

Barron, B. (2003). When smart groups fail. *Journal of the Learning Sciences, 12*(3), 307-359.

Berland, L. K., & Reiser, B. J. (2009). Making sense of argumentation and explanation. *Science Education, 93*(1), 26-55.

Betts, B. (2013, March 12). Gamification, meet gamefulness. *Learning Technologies Blog*. Retrieved from http://www.astd.org/Publications/Blogs/Learning-Technologies-Blog/2013/03/Gamification-Meet-Gamefulness

Bloom, B. S. (1984). *Taxonomy of educational objectives book 1: Cognitive domain* (2nd ed.). New York: Addison Wesley.

Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., & Schulz, L. (2011). The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition, 120*(3), 322-330. doi: 10.1016/j.cognition.2010.10.001

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy.

Bressler, D. M., & Bodzin, A. M. (2013). A mixed methods assessment of students' flow experience during a mobile augmented reality science game. *Journal of Computer Assisted Learning, 29*(6), 505-517.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32-42.

Brown, S. (2009). *Play: How it shapes the brain, opens the imagination, and invigorates the soul*. New York: Avery.

Bruner, J. S. (1961). The act of discovery. *Harvard Educational Review, 31*(1), 21-32.

Bruner. J. S. (1966). *Toward a theory of instruction*. Cambridge, MA: The Belknap Press.

Bruner, J. (1983). Play, thought, and language. *Peabody Journal of Education, 60*(3), pp. 60-69.

Cabrera, J. S., Frutos, H. M., Stoica, A. G., Avouris, N., Dimitriadis, Y., Fiotakis, G., & Liveri, K. D. (2005). *Mystery in the museum: Collaborative learning activities using handheld devices*. Paper presented at the 7th International Conference on Human-Computer Interaction with Mobile Devices, Salzburg, Austria.

Casner-Lotto, J., Rosenblum, E., & Wright, M. (2009). *The ill-prepared workforce: Exploring the challenges of employer-provided workforce readiness training*. New York: The Conference Board. Retrieved from http://www.cvworkingfamilies.org/system/files/Ill_preparedWorkforce_RR.pdf

Chatterjee, S., Mohanty, A., & Bhattacharya, B. (2011). Peer collaboration, facilitator intervention, and learning styles in computer game-based learning: Initial findings from an empirical study. *Proceedings of the European Conference on Games Based Learning*, 683-690.

Chang, C.-S., Chen, T.-S., & Hsu, W.-H. (2011). The study on integrating WebQuest with mobile learning for environmental education. *Computers & Education, 57*(1), 1228-1239.

Cheng, K.-H., & Tsai, C.-C. (2012). Affordances of augmented reality in science learning: Suggestions for future research. *Journal of Science Education and Technology* (3 August 2012), 1-14.

Christensen, C. M., Horn, M. B., & Johnson, C. W. (2008). *Disrupting class: How disruptive innovation will change the way the world learns*. New York: McGraw-Hill.

Clark, R. E. (2007). Learning from serious games? Arguments, evidence, and research suggestions. *Educational Technology, 47*(3), 56-59.

Clark, D. B., & Martinez-Garza, M. (2012). Prediction and explanation as design mechanics in conceptually integrated digital games to help players articulate the tacit understandings they build through game play. In C. Steinkuehler, K. Squire, & S. Barab (Eds.), *Games, learning, and society: Learning and meaning in the digital age* (pp. 279-305). New York: Cambridge University Press.

Clark, D. B., Nelson, B., Sengupta, P., & D'Angelo, C. M. (2009). Rethinking science learning through digital games and simulations: Genres, examples, and evidence *Invited Topic Paper in the Proceedings of The National Academies Board on Science Education Workshop on Learning Science: Computer Games, Simulations, and Education.* Washington, DC.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.

Coulter, B., Klopfer, E., Sheldon, J., & Perry, J. (2012). Discovering familiar places: Learning through mobile place-based games. In C. Steinkuehler, K. Squire, & S. Barab (Eds.), *Games, learning, and society: Learning and meaning in the digital age*. New York: Cambridge University Press.

Crawford, C. (2011). *The art of computer game design* [Kindle version]. Retrieved from Amazon.com. (Original work published by 1984)

Creswell, J., & Plano Clark, V. (2011). *Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Crompton, H. (2013). Mobile learning: New approach, new theory. In Z. Berge & L. Muilenburg (Eds.). *Handbook of Mobile Learning* (pp. 47-57). New York: Routledge.

Csikszentmihalyi, M. (1975). Play and intrinsic rewards. *Journal of Humanistic Psychology, 15*(3), 41-63.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*: New York: Harper and Row.

Csikszentmihalyi, M. (1996). *Creativity: Flow and the psychology of discovery and invention.* New York: Harper Collins.

Csikszentmihalyi, M. (1997*). Finding flow: The psychology of engagement with everyday life.* New York: Harper Collins.

Csikszentmihalyi, M., Rathunde, K., & Whalen, S. (1993). *Talented teenagers: The roots of success and failure*: New York: Cambridge University Press.

Cuban, L. (2001). *Oversold and underused: Computers in the classroom*. Cambridge, MA: Harvard University Press.

Czarnecki, K. (2012). Mobile gaming in public libraries. In S. Dikkers, J. Martin, & B. Coulter (Eds.), *Mobile media learning: Amazing uses of mobile devices for teaching and learning* (pp. 131-140). Pittsburg, PA: ETC Press.

Davidson, C. N., & Goldberg, D. T. (2010). *The future of thinking: Learning institutions in a digital age*. Cambridge, MA: The MIT Press.

Dede, C. (2009). Immersive interfaces for engagement and learning. *Science, 323*(5910), 66-69.

Dede, C., Grotzer, T. A., Metcalf, S. J., & Kamarainen, A. M. (2011). *EcoMOBILE: Blended learning across virtual and natural ecosystems*. [Web page]. Retrieved from http://ecomobile.gse.harvard.edu/

Demetriadis, S., Tsiatsos, T., & Karakostas, A. (2012). Scripted collaboration to guide the pedagogy and architecture of digital learning games. *Proceedings of the European Conference on Games Based Learning*, 148-154.

Desy, E. A., Peterson, S. A., & Brockman, V. (2011). Gender differences in science-related attitudes and interests among middle school and high school students. *Science Educator, 20*(2), 23-30.

Dewey, J. (1997). *How we think*. Mineola, NY: Dover Publications. (Original work published 1910).

Dieterle, E. (2009). Neomillennial learning styles and River City. *Children, Youth & Environments, 19*(1), 245-278.

Dieterle, E., Dede, C., & Schrier, K. (2007). "Neomillennial" learning styles propagated by wireless handheld devices. In M. Lytras & A. Naeve (Eds.), *Ubiquitous and pervasive* *knowledge and learning management: Semantics, social networking and new media to their full potential*. Hershey, PA: Idea Group, Inc.

Dikkers, S. (2012). History in our hands: Mobile media in museum adventures. In S. Dikkers, J. Martin, & B. Coulter (Eds.), *Mobile media learning: Amazing uses of mobile devices for teaching and learning* (pp. 121-130). Pittsburg, PA: ETC Press.

Dillenbourg, P. (1999). What do you mean by "collaborative learning"? In P. Dillenbourg (Ed.), *Collaborative learning: Cognitive and computational approaches* (pp. 1-16). Amsterdam: Pergamon, Elsevier Science.

Dovis, S., Oord, S., Wiers, R., & Prins, P. (2012). Can motivation normalize working memory and task persistence in children with attention-deficit/hyperactivity disorder? The effects of money and computer-gaming. *Journal of Abnormal Child Psychology, 40*(5), 669-681. doi: 10.1007/s10802-011-9601-8

Driscoll, M. P. (2005). *Psychology of learning for instruction* (3rd ed.). Boston, MA: Pearson Education.

Dunleavy, M. & Dede, C. (2013). Augmented reality teaching and learning. In J.M. Spector, M.D. Merrill, J. Elen & M.J. Bishop (Eds.). *Handbook of Research on Educational Communications and Technology*: *Volume 2 (*4th ed.). New York: Springer.

Dunleavy, M., Dede, C., & Mitchell, R. (2009). Affordances and limitations of immersive participatory augmented reality simulations for teaching and learning. *Journal of Science Education and Technology, 18*(1), 7–22.

Dunleavy, M. & Simmons, B. (2011). Assessing learning and identity in augmented reality science games. In L. Annetta & S. Bronack (Eds.), *Serious educational games assessment* (pp. 221-240). Rotterdam, The Netherlands: Sense Publishers.

Dwyer, D. (1994). Apple Classrooms of Tomorrow: What we've learned. *Educational Leadership, 51*(7), 4-10.

Dynarski, M., Agodini, R., Heaviside, S., Novak, T., Carey, N., Campuzano, L...Sussex, W. (2007). *Effectiveness of Reading and Mathematics Software Products: Findings from the First Student Cohort.* Washington, D.C.: U.S. Department of Education, Institute of Education Sciences.

Ellis, C. R. (2007). No Child Left Behind--A critical analysis. *Curriculum & Teaching Dialogue, 9*(1/2), 221-233.

Facer, K., Joiner, R., Stanton, D., Reid, J., Hull, R., & Kirk, D. (2004). Savannah: Mobile gaming and learning? *Journal of Computer Assisted Learning, 20*(6), 399-409.

Faiola, A., Newlon, C., Pfaff, M., & Smyslova, O. (2013). Correlating the effects of flow and telepresence in virtual worlds: Enhancing our understanding of user behavior in game-based learning. *Computers in Human Behavior, 29*(3), 1113-1121.

Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research, 59*(2), 117-142.

Fisher, C., Dwyer, D. C., & Yocam, K. (Eds.) (1996). *Education and technology: Reflections on computing in classrooms*. San Francisco: Jossey-Bass Publishers.

Folta, E. E. (2010). *Investigating the impact on student learning and outdoor science interest through modular Serious Educational Games: A design-based research study*. (Unpublished doctoral dissertation). North Carolina State University, Raleigh, NC.

Foreman, J. (2004). Game based learning: How to delight and instruct in the 21st century. *EDUCAUSE Review*, 39(5), 50–66.

Fotouhi-Ghazvini, F., Earnshaw, R. A., Robison, D., & Excell, P. S. (2009, September). *Designing augmented reality games for mobile learning using an instructional-*

motivational paradigm. Paper presented at the International Conference on CyberWorlds, Bradford, UK.

Gatto, J. T. (2009). *Weapons for mass instruction: A schoolteacher's journey through the dark world of compulsory schooling*. Gabriola Island, BC, Canada: New Society Publishers.

Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. New York: Palgrave Macmillan.

Gee, J. P. (2007). *Good video games + good learning: Collected essays on video games, learning, and literacy*. New York: Peter Lang Publishing.

Gee, J. P. (2010). *New digital media and learning as an emerging area and "worked examples" as one way forward*. Cambridge, MA: The MIT Press.

Gladwell, M. (2000). *The tipping point: How little things can make a big difference*. New York: Little, Brown and Company.

Gopnick, A., Meltzoff, A. N., & Kuhl, P. K. (1999). *The scientist in the crib: What early learning tells us about the mind*. New York: HarperCollins Publishers.

Goyal, N. (2012). *One size does not fit all: A student's assessment of school* [Kindle version]. Retrieved from Amazon.com

Grimley, M., Green, R., Nilsen, T., Thompson, D., & Tomes, R. (2011). Using computer games for instruction: The student experience. *Active Learning in Higher Education, 12*(1), 45-56.

Gustafsson, A., Katzeff, C., & Bang, M. (2009). Evaluation of a pervasive game for domestic energy engagement among teenagers. *ACM Computers in Entertainment, 7*(4), 54-72.

Hatfield, D. (2011). *The right kind of telling: An analysis of feedback and learning in a journalism epistemic game.* (Unpublished doctoral dissertation). University of Wisconsin-Madison, Madison, WI.

Healy, J. M. (1998). *Failure to connect: How computers affect our children's minds—and what we can do about it*. New York: Simon & Schuster.

Hee-Sun, L., Liu, O. L., & Linn, M. C. (2011). Validating measurement of knowledge integration in science using multiple-choice and explanation items. *Applied Measurement in Education, 24*(2), 115-136.

Heneghan, J. (2008). Foreword. In L. A. Annetta, *Serious educational games: From theory to practice* (pp. ix-xi). Rotterdam, The Netherlands: Sense Publishers.

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

Honey, M., & Henriquez, A. (1993). *Telecommunications and K-12 educators: Findings from a national survey*. New York: Center for Technology in Education.

Howell, D. D. (2010). The riddle of the smart machines. *TechTrends: Linking Research & Practice to Improve Learning, 54*(1), 33-37.

Huang, W.-H. (2011). Evaluating learners' motivational and cognitive processing in an online game-based learning environment. *Computers in Human Behavior, 27*(2), 694-704.

Huizinga, J. (1971). *Homo ludens: A study of the play-element in culture* (J. R. F. C. Hull, Trans.). Boston: The Beacon Press. (Original work published 1938)

IBM Corp. (2012). *IBM SPSS Statistics for Windows* (Version 21.0) [Computer software]. Armonk, NY: Author.

Inal, Y., & Cagiltay, K. (2007). Flow experiences of children in an interactive social game environment. *British Journal of Educational Technology, 38*(3), 455-464.

Ito, M., Baumer, S., Bittanti, M., boyd, d., Cody, R., Herr-Stephenson, B., . . . Tripp, L. (2010). *Hanging out, messing around, geeking out: Kids living and learning with new media*. Cambridge, MA: MIT Press.

Jackson, S., Eklund, R., & Martin, A. (2010). *The FLOW manual*. Queensland, Australia: Mind Garden, Inc.

Johnson, L., Adams, S., & Cummins, M. (2012). *NMC horizon report: 2012 K-12 edition*. Austin, Texas: The New Media Consortium.

Johnson, L., Adams Becker, S., Cummins, M., Estrada V., Freeman, A., and Ludgate, H. (2013). *NMC Horizon Report: 2013 K-12 Edition.* Austin, Texas: The New Media Consortium.

Johnson, D. W., & Johnson, R. T. (1989). *Cooperation and competition: Theory and research*. Edina, Minnesota: Interaction Book Company.

Johnson, D. W., & Johnson, R. T. (1999). Making cooperative learning work. *Theory Into Practice, 38*(2), 67-73.

Johnson, D. W., Johnson, R. T., & Holubec, E. J. (1993). *Circles of learning: Cooperation in the classroom* (4th ed.). Edina, Minnesota: Interaction Book Company.

Kamarainen, A. M., Metcalf, S. J., Grotzer, T. A., & Dede, C. (2012, June). *EcoMOBILE: Blended learning across virtual and natural ecosystems*. Poster session presented at the PI Meeting for Discovery Research K-12, Hyatt Regency Crystal City, VA. Retrieved from http://cadrek12.org/sites/default/files/Dede%20Ecomobile%20poster.jpg

Kanter, R. M. (2001). *Evolve! Succeeding in the digital culture of tomorrow*. Cambridge, MA: Harvard Business School Press.

Katz, J. (2006). *Magic in the air: Mobile communication and the transformation of social life*. New Brunswick, NJ: Transaction Publishers.

Ketelhut, D. (2007). The impact of student self-efficacy on scientific inquiry skills: An exploratory investigation in River City, a multi-user virtual environment. *Journal of Science Education & Technology, 16*(1), 99-111.

Ketelhut, D. J., & Nelson, B. C. (2010). Designing for real-world scientific inquiry in virtual environments. *Educational Research, 52*(2), 151-167.

Ketelhut, D. J., Nelson, B. C., Clarke, J., & Dede, C. (2010). A multi-user virtual environment for building and assessing higher order inquiry skills in science. *British Journal of Educational Technology, 41*(1), 56-68.

Khatib, F., DiMaio, F., Cooper, S., Kazmierczyk, M., Gilski, M., Krzywda, S., . . . Baker, D. (2011). Crystal structure of a monomeric retroviral protease solved by protein folding game players. *Nature Structural & Molecular Biology, 18*(10), 1175-1177.

Kiili, K., & Lainema, T. (2008). Foundation for measuring engagement in educational games. *Journal of Interactive Learning Research, 19*(3), 469-488.

Kirriemuir, J., & McFarlane, A. (2004). *Report 8: Literature review in games and learning*. Bristol, UK: NESTA Futurelab.

Kirshner, D. & Whitson, J. A. (1997). Editor's introduction. In D. Kirshner & J. A. Whitson (Eds.), *Situated cognition: Social, semiotic, and psychological perspectives*. Mahwah, NJ: Erlbaum.

Klopfer, E. (2008). *Augmented learning: Research and design of mobile educational games*. Cambridge, MA: The MIT Press.

Klopfer, E., Osterweil, S., & Salen, K. (2009). *Moving learning games forward.* Cambridge, MA: Education Arcade. Retrieved from http://education.mit.edu/papers/MovingLearningGamesForward_EdArcade.pdf

Klopfer, E., Perry, J., Squire, K., & Jan, M.-F. (2005). *Collaborative learning through augmented reality role playing*. Paper presented at the meeting of Computer Supported Collaborative Learning, Taipei, Taiwan.

Klopfer, E., Perry, J., Squire, K., Jan, M.-F., & Steinkuehler, C. (2005). *Mystery at the museum: A collaborative game for museum education*. Paper presented at the annual conference of Computer Supported Collaborative Learning, Taipei, Taiwan.

Klopfer, E., & Sheldon, J. (2010). Augmenting your own reality: Student authoring of science-based augmented reality games. *New Directions for Youth Development, 2010*(128), 85-94.

Klopfer, E., Sheldon, J., Perry, J., & Chen, V. H. H. (2012). Ubiquitous games for learning (UbiqGames): Weatherlings, a worked example. *Journal of Computer Assisted Learning, 28*(5), 465-476.

Klopfer, E., & Squire, K. (2008). Environmental Detectives—the development of an augmented reality platform for environmental simulations. *Educational Technology Research & Development, 56*(2), 203-228.

Klopfer, E., Squire, K., & Jenkins, H. (2002). *Environmental detectives: PDAs as a window into a virtual simulated world*. Paper presented at the International Conference on Wireless Mobile Technologies in Education, Växjö, Sweden.

Kohn, A. (2000). *The case against standardized testing: Raising the scores, ruining the schools.* Portsmouth, NH: Heinemann.

Lai, C. Y., & Wu, C. C. (2006). Using handhelds in a jigsaw cooperative learning environment. *Journal of Computer Assisted Learning, 22*(4), 284-297.

Laurillard, D. (2009). Foreword. In G. Vavoula, N. Pachler, & A. Kukulska-Hulme (Eds.), *Researching mobile learning: Frameworks, tools and research designs* (pp. ix-xi). Bern, Switzerland: Peter Lang.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York: Cambridge University Press.

Leech, N. L, Barrett, K. C., & Morgan, G. A. (2008). *SPSS for intermediate statistics: Use and interpretation* (3rd ed.). New York: Lawrence Erlbaum Associates.

Lenhart, A., Hitlin, P., & Madden, M. (2005). *Teens and technology*. Washington, DC: Pew Internet & American Life Project. Retrieved from http://pewinternet.org/Reports/2005/Teens-and-Technology.aspx

Leonard, D. (2013, October 24). The iPad goes to school, *Business Week*. Retrieved February 11, 2014 from http://www.businessweek.com/printer/articles/162786-the-ipad-goes-to-school

Lincoln, Y., & Guba, E. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage Publications, Inc.

Lindahl, B. (2007, April). *A longitudinal study of students' attitudes towards science and choice of career.* Paper presented at the National Association for Research in Science Teaching, New Orleans.

Lockard, C. B., & Wolf, M. (2012). *Occupational employment projections to 2020*. Washington, DC: U.S. Bureau of Labor Statistics. Retrieved from http://www.bls.gov/opub/mlr/2012/01/art5full.pdf

Lomax, R. G. (2001). *Statistical concepts: A second course for education and the behavioral sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Looi, C. K., Zhang, B., Chen, W., Seow, P., Chia, G., Norris, C., & Soloway, E. (2011). 1:1 mobile inquiry learning experience for primary science students: A study of learning effectiveness. *Journal of Computer Assisted Learning, 27*(3), 269-287.

Malone, T. W. (1981). Toward a theory of intrinsically motivating instruction. *Cognitive Science, 4*, 333–369.

Mansour, S. S., & El-Said, M. (2009). Multi-players role-playing educational serious games: A link between fun and learning. *International Journal of Learning, 15*(11), 229-239.

Mathews, J., Holden, C., Jan, M.-F., & Martin, J. (2008). Sick at South Shore Beach: A place-based augmented reality game as a framework for building evidence-based arguments. *Proceedings of 8th International Conference on Learning Sciences 3*, 89-90.

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

Medina, J. (2008). *Brain rules: 12 principles for surviving and thriving at work, home, and school*. Seattle, WA: Pear Press.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Montessori, M. (1912). *The Montessori method* (A.E. George, Trans.). New York: Frederick A. Stokes Company.

Mooney, C.G. (2000). *Theories of childhood: An introduction to Dewey, Montessori, Erikson, Piaget, & Vygotsky*. St. Paul, MN: Redleaf Press.

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Washington, DC: U.S. Government Printing Office.

National Research Council. (1995). *Reinventing schools: The technology is now!* Washington, DC: The National Academies Press.

National Research Council. (2007a). *Rising above the gathering storm: Energizing and employing America for a brighter economic future*. Washington, DC: The National Academies Press.

National Research Council. (2007b). *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: The National Academies Press.

National Research Council. (2011a). *Learning science through computer games and simulations*. Washington, DC: The National Academies Press.

National Research Council. (2011b). *High school dropout, graduation, and completion rates: Better data, better measures, better decisions*. Washington, DC: The National Academies Press.

National Research Council. (2012a). *A framework for K-12 science education: Practices, cross-cutting concepts, and core ideas*. Washington, DC: The National Academies Press.

National Research Council. (2012b). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. Washington, DC: The National Academies Press.

Negroponte, N. (1995). *Being digital*. New York: Alfred A. Knopf.

Nelson, B. C., & Ketelhut, D. J. (2007). Scientific inquiry in educational multi-user virtual environments. *Educational Psychology Review, 19*(3), 265-283.

Norris, C. A., & Soloway, E. (2011). Learning and schooling in the age of mobilism. *Educational Technology, 51*(6), 3-12.

Nussbaum, E. M. (2012). Argumentation and student-centered learning environements. In D. Jonassen & S. Land (Eds.), *Theoretical foundations of learning environments: Second edition* (pp. 114-141). New York: Routledge.

O'Shea, P., Dede, C., & Cherian, M. (2011). The results of formatively evaluating an augmented reality curriculum based on modified design principles. *International Journal of Gaming and Computer-Mediated Simulations, 3*(2), 57-66.

Palincsar, A. S. (1998). Social constructivist perspectives on teaching and learning. *Annual Review of Psychology, 49*(1), 345-375.

Papert, S. (1996). *The connected family: Bridging the digital generation gap*. Marietta, GA: Longstreet Press.

Papert, S. (2002). Hard fun. *Bangor News Daily*. Retrieved from http://www.papert.org/articles/HardFun.html

Park, J., Parsons, D., & Ryu, H. (2010). To flow and not to freeze: Applying flow experience to mobile learning. *IEEE Transactions on Learning Technologies, 3*(1), 56-67.

Partnership for 21st Century Skills. (2013). *Overview* [Web page]. Retrieved from http://www.p21.org/overview

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Pentland, A. (2012). The new science of building great teams. *Harvard Business Review, 90*(4), 60-69.

Perry, J., Klopfer, E., Norton, M., Sutch, D., Sandford, R., & Facer, K. (2008). *AR gone wild: two approaches to using augmented reality learning games in Zoos*. Paper presented at the 8th international for the learning sciences, Utrecht, Netherlands.

Piaget, J. (1962). *Play, dreams, and imitations in childhood* (C. Gattegno & F. M. Hodgson, Trans.). New York: W.W. Norton and Company.

Piaget, J. (1985). *The equilibration of cognitive structures: The central problem of intellectual development* (T. Brown & K. J. Thampy, Trans.). Chicago: University of Chicago Press.

Prensky, M. (2001). *Digital game-based learning*. [Kindle version]. Retrieved from Amazon.com

Project Tomorrow. (2010). *Creating our future: Students speak up about their vision for 21st century learning.* Irvine, CA: Project Tomorrow. Retrieved from http://www.tomorrow.org/speakup/pdfs/SU09NationalFindingsStudents&Parents.pdf

Powell, W. W., & Snellman, K. (2004). The knowledge economy. *Annual Review of Sociology, 30*, 199–220.

Ramirez, A. (2013). *Save our science: How to inspire a new generation of scientists* [Kindle version]. Retrieved from Amazon.com

Randel, J. M., & Morris, B. A. (1992). The effectiveness of games for educational purposes: A review of recent research. *Simulation & Gaming, 23*(3), 261-276.

Rathunde, K., & Csikszentmihalyi, M. (2005). The social context of middle school: Teachers, friends, and activities in Montessori and traditional school environments. *Elementary School Journal, 106*(1), 59-79.

Ravitch, D. (2011). *The death and life of the great American school system: How testing and choice are undermining education* [Kindle version]. Retrieved from Amazon.com

Reigeluth, C. M. (1994). Introduction: The imperative for systemic change. In C. M. Reigeluth & R. J. Garfinkle (Eds.), *Systemic change in education* (pp. 3-14). Englewood Cliffs, NJ: Educational Technology Publications, Inc.

Reiser, B. J., Berland, L. K., & Kenyon, L. (2012). Engaging students in the scientific practices of explanation and argumentation. *Science Teacher, 79*(4), 34-39.

Richardson, W. (2012). *Why school? How education must change when learning and information are everywhere*. [Kindle version]. Retrieved from Amazon.com

Roschelle, J., & Pea, R. (2002). A walk on the WILD side: How wireless handhelds may change computer-supported collaborative learning. *International Journal of Cognition and Technology*, 1(1), 145-168.

Rosenbaum, E., Klopfer, E., & Perry, J. (2007). On location learning: Authentic applied science with networked augmented realities. *Journal of Science Education & Technology, 16*(1), 31-45.

Rugg, H., & Shumaker, A. (1928). *The child-centered school*. New York: Arno & The New York Times.

Ryu, H., & Parsons, D. (2012). Risky business or sharing the load? Social flow in collaborative mobile learning. *Computers & Education, 58*(2), 707-720.

Sadler, T. D., Romine, W. L., Stuart, P. E., & Merle-Johnson, D. (2013). Game-based curricula in biology classes: Differential effects among varying academic levels. *Journal of Research in Science Teaching, 50*(4), 479-499.

Salen, K., Torres, R., Wolozin, L., Rufo-Tepper, R., & Shapiro, A. (2011). *Quest to Learn: Developing the school for digital kids*. Cambridge, MA: The MIT Press.

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

Sandberg, S. (2013). *Lean in: Women, work, and the will to lead*. New York: Random House.

Sánchez, J., & Olivares, R. (2011). Problem solving and collaboration using mobile serious games. *Computers & Education, 57*(3), 1943-1952.

Schrage, M. (2000). *Serious play: How the world's best companies simulate to innovate*. Boston: Harvard Business School Press.

Schwabe, G., & Göth, C. (2005). Mobile learning with a mobile game: Design and motivational effects. *Journal of Computer Assisted Learning, 21*(3), 204-216.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental design and quasi-experimental designs for generalized causal inferences*. Boston: Houghton Mifflin.

Shaffer, D. W. (2006). *How computer games help children learn*. New York: Palgrave Macmillan.

Shaffer, D. W., Hatfield, D., Svarovsky, G. N., Nash, P., Nulty, A., Bagley, E., . . . Mislevy, R. (2009). Epistemic Network Analysis: A prototype for 21st century assessment of learning. *International Journal of Learning and Media, 1*(2), 1-21.

Sharritt, M. J. (2008). Forms of learning in collaborative video game play. *Research & Practice in Technology Enhanced Learning, 3*(2), 97-138.

Shelton, B. E. (2008). Is educational game research doomed to fail? *TechTrends: Linking Research & Practice to Improve Learning, 52*, 74-75.

Singh, K., Mido, C., & Dika, S. (2005). Affective and motivational factors in engagement and achievement in science. *International Journal of Learning, 12*(6), 207-218.

Sitzmann, T. (2011). A meta-analytic examination of the instructional effectiveness of computer-based simulation games. *Personnel Psychology, 64*(2), 489-528.

Sivin-Kachala, J., & Bialo, E. R. (1996). *The effectiveness of technology in schools: A summary of recent research*. New York: Interactive Educational Systems Design.

Somervill, B. A. (2010). *Great empires of the past: Empires of Ancient Mesopotamia*. New York: Chelsea House.

Songer, N. B. (1996). Exploring learning opportunities in coordinated network-enhanced classrooms: A case of kids as global scientists. *Journal of the Learning Sciences, 5*(4), 297-327.

Spikol, D., & Milrad, M. (2008). Physical activities and playful learning using mobile games. *Research & Practice in Technology Enhanced Learning, 3*(3), 275-295.

Spires, H. A., Lee, J. K., Turner, K. A., & Johnson, J. (2008). Having our say: Middle grade student perspectives on school, technologies, and academic engagement. *Journal of Research on Technology in Education, 40*(4), 497-515.

Squire, K. D. (2004). *Replaying history: Learning world history through playing Civilization III*. (Unpublished doctoral dissertation). Indiana University, Bloomington, IN.

Squire, K. (2005). Changing the game: What happens when video games enter the classroom? *Innovate, 1*(6). Retrieved from http://www.innovateonline.info/index.php?view=article&id=82

Squire, K. (2006). From content to context: Videogames as designed experience. *Educational Researcher*, *35*(8), 19–29.

Squire, K. (2010). From information to experience: Place-based augmented reality games as a model for learning in a globally networked society. *Teachers College Record, 112*(10), 2565-2602.

Squire, K. (2011). *Video games and learning: Teaching and participatory culture in the digital age*. New York: Teachers College Press.

Squire, K., & Jan, M. (2007). Mad city mystery: Developing scientific argumentation skills with a place-based augmented reality game on handheld computers. *Journal of Science Education & Technology, 16*(1), 5-29.

Squire, K., & Klopfer, E. (2007). Augmented reality simulations on handheld computers. *Journal of the Learning Sciences, 16*(3), 371-413.

Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-supported collaborative learning. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 409-425). New York: Cambridge University Press.

Steinkuehler, C. (2006). Massively multiplayer online videogaming as participation in a discourse. *Mind, Culture, and Activity, 13*(1), 38–52.

Steinkuehler, C., & Duncan, S. (2008). Scientific habits of mind in virtual worlds. *Journal of Science Education & Technology, 17*(6), 530-543.

Stevens, R. J., & Slavin, R. E. (1995). The cooperative elementary school: Effects on students' achievement, attitudes, and social relations. *American Educational Research Journal*, 32, 321–351.

SurveyMonkey. (1999-2014). *SurveyMonkey.* Retrieved from http://www.surveymonkey.com/

Tai, R. H., Liu, C. Q., Maltese, A. V., & Fan, X. (2006). Planning early for careers in science. *Science, 312*(5777), 1143-1144.

Tapscott, D. (2009). *Grown up digital: How the net generation is changing your world*. New York: McGraw-Hill.

Tapscott, D., & Williams, A. D. (2006). *Wikinomics: How mass collaboration changes everything.* New York: Penguin Group.

Thomas, D., & Brown, J. S. (2011). *A new culture of learning: Cultivating the imagination for a world of constant change*. Lexington, Ky: CreateSpace Independent Publishing Platform.

Trespalacios, J., Chamberlin, B., & Gallagher, R. (2011). Collaboration, engagement & fun: How youth preferences in video gaming can inform 21st century education. *TechTrends: Linking Research & Practice to Improve Learning, 55*(6), 49-54.

Vogel, J. J., Vogel, D. S., Cannon-Bowers, J. A. N., 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.

Vygotsky, L. S. (1966). Play and its role in the mental development of the child. *Soviet Psychology, 12*(6), 62-76.

Vygotsky, L.S. (1978). *Mind in society: Development of higher psychological processes* (14th ed.). Cambridge, MA: Harvard University Press.

Wagler, M., & Mathews, J. (2012). Up River: Place, ethnography, and design in the St. Louis River estuary. In S. Dikkers, J. Martin, & B. Coulter (Eds.), *Mobile media learning: Amazing uses of mobile devices for teaching and learning* (pp. 23-40). Pittsburg, PA: ETC Press.

Wang, C. K. J., Khoo, A., Liu, W. C., & Divaharan, S. (2008). Passion and intrinsic motivation in digital gaming. *CyberPsychology & Behavior, 11*(1), 39–45.

Walker, C. J. (2010). Experiencing flow: Is doing it together better than doing it alone? *The Journal of Positive Psychology, 5*(1), 3-11.

Watkins, J., & Mazur, E. (2013). Retaining students in Science, Technology, Engineering, and Mathematics (STEM) majors. *Journal of College Science Teaching, 42*(5), 36-41.

Watson, W. R. (2010). Games for social studies. In A. Hirumi (Ed.), *Playing games in school: Video games and simulations for primary and secondary education* (pp. 173-202). Washington, DC: International Society for Technology in Education.

Wenglinsky, W. (2005). *Using technology wisely: The keys to success in schools*. New York: Teachers College Press.

Willingham, D. T. (2010). *Why don't students like school: A cognitive scientist answers questions about how the mind works and what it means for the classroom*. San Francisco, CA: Jossey-Bass.

Wood, R. T. A., Griffiths, M. D., & Parke, A. (2007). Experiences of time loss among videogame players: An empirical study. *CyberPsychology and Behavior, 10*(1), 38-44.

Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., & Malone, T. W. (2010). Evidence for a collective intelligence factor in the performance of human groups. *Science, 330*(6004), 686-688. doi: 10.1126/science.1193147

Woolley, A. W., Gerbasi, M. E., Chabris, C. F., Kosslyn, S. M., & Hackman, J. R. (2008). Bringing in the experts: How team composition and collaborative planning jointly shape analytic effectiveness. *Small Group Research, 39*(3), 352-371.

Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, Advance online publication. doi: 10.1037/a0031311

Wrzesien, M., & Alcañiz Raya, M. (2010). Learning in serious virtual worlds: Evaluation of learning effectiveness and appeal to students in the E-Junior project. *Computers & Education, 55*(1), 178-187.

Yin, R. K. (2014). *Case study research: Design and methods* (5th ed.). Thousand Oaks, CA: Sage Publications, Inc.

Young, M. F., Slota, S., Cutter, A. B., Jalette, G., Mullin, G., Lai, B., . . . Yukhymenko, M. (2012). Our princess is in another castle: A review of trends in serious gaming for education. *Review of Educational Research, 82*(1), 61-89.

Appendix Index

APPENDIX A: Game Overview

School Scene Investigators: The Case of the Mystery Powder was an augmented reality mobile learning game where students investigated a mysterious powder at their school. Someone has stolen money from the cash register in the cafeteria and left behind this unknown white powder. Was it the janitor who used cornstarch as a cleaning agent? Was it the front office secretary who baked sugar cookies for the staff? Or was it the student who was making a rocket using baking soda? All had valid reasons to get supplies from the kitchen, but who stole the cash?

The game took place as five chapters, roughly aligning to one chapter per class period. Chapter #1: Students were introduced to the incident and the main characters. Students visited the principal and he asked for their help with the investigation; he indicated that there would be a detective on-site who might be of assistance. Then, they visited the cafeteria to explore the crime scene. Finally, they met and discussed the incident with the three main suspects: the janitor, the secretary, and the fictional fellow student. They also had a brief conversation with their science teacher in order to gain some helpful scientific knowledge. Chapter #2: Students visited areas of the school where the suspects left evidence. At each location, they found evidence of the known powders and conducted some simple, virtual tests including vinegar, iodine, heat, and pH tests. Content knowledge and tests results were all conveyed using pictures and videos during gameplay. There was also a brief encounter with the principal which served as a review of what was learned so far in the game. At the conclusion of Chapter #2, players synthesized the evidence gathered so far by completing part of an incident report. Chapter #3: The detective visited the students and brought them a sample of the mystery powder that had been cleared by his lab for student testing. Facilitated by some teacher instruction and assistance, student teams then conducted their own tests on an actual mystery powder. The detective supplied helpful information for what to look for as they conducted each test, see Figure A1.

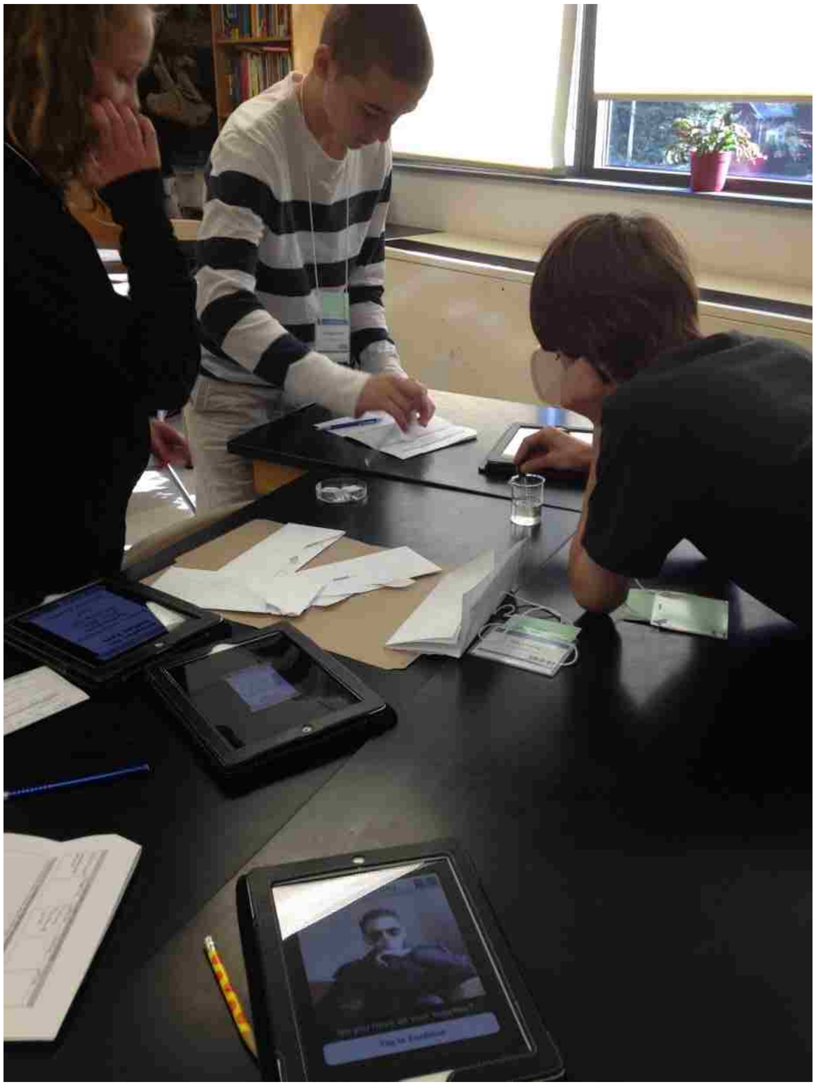


Figure A1: Players using iPads for experiment instructions.

At the conclusion of Chapter #3, students analyzed the new information about the mystery powder, wrote up a final conclusion about what powders were present in the powder, and then narrowed down the suspect list to two people. Chapter #4: Since there were still two possible suspects, the detective recommended revisiting the crime scene to see if they missed anything. Sure enough, upon arrival back at the cafeteria, one of the players noticed a tool mark on the cash register—a tool was used to force open the cash register! Teams then revisited the locations where suspects stored their belongings and found out what tools each suspect had access to. They took the tools virtually with them to the art room and made clay impressions to compare the tool mark from the crime scene to the suspects' tools. Chapter #5: Once students decided who the real thief was they revisited the principal to give their final accusation.

During gameplay, students utilized the scientific method, learned about acids and bases, and conducted some basic physical and chemical tests to analyze evidence and solve the mystery. This narrative-driven, inquiry-based game was played on iPads supplied by the Middle School. As students moved throughout their school building, they encountered quick-response codes (QR codes) that they scanned to access information relevant to the game. This included conversing with a virtual character and the retrieval of evidence to keep in their inventory. Players were also required to talk to real people in the building to get additional game information. Players also deciphered a code and typed in the answer manually to the decoder.

The game was played in teams of three or four where each student had a unique role: social networker, techie, photographer, or pyro-technician. Based on their role, they were provided with different pieces of information as they progressed through the game. In order to effectively collaborate and solve the mystery, players had to share their information and work together.

In order to complete the game, students utilized next generation science skills such as asking questions, carrying out an investigation, data analysis, constructing explanations, and engaging in argument from evidence. Throughout the game, players were exposed to some basic forensic science, like analyzing tool marks. They also solved other problems along the way such as decoding a locker combination and determining whether the characters they met had a motive.

APPENDIX B: Original Documentation for Mystery Powder Lab Activity

Mystery Mixture Lab

Objective: to determine the compounds in an unknown mixture by using experimentation.

Hypothesis: If ____
____
Then ____
____.

Materials:

test tubes (4)	test tube holder	vinegar
test tube rack	cornstarch	baking soda
sugar	mystery mixture	petri Dish
eyedroppers (2)	iodine solution	water

Procedure:

1. Place a scoopful (**One at a time)** of cornstarch, baking soda, and sugar in a **Petri Dish**. *You must leave some of the mixture in the test tube for heating*
2. Add three (3) drops of vinegar to each substance.
3. Record observations.
4. Wash and dry the ***Petri dish***.
5. Place a scoopful of cornstarch, baking soda, and sugar in the ***Petri Dish***.
6. Add three (3) drops of iodine solution to each substance.
7. Record observations.
8. Wash and dry the ***Petri Dish***.
9. Place a scoopful of cornstarch, baking soda, and sugar in the ***Petri Dish***.
10. Add three drops of water to each substance. Complete Litmus paper test. record your results on chart paper.
11. Take the remaining amount of each substance in the test tube. (Cornstarch in one, sugar in another etc.)
12. Gently heat the bottom of each test tube. Do not put the test tube directly in the flame.
13. Record observations.
14. Pour mystery mixture into the ***Petri Dish***. Be sure to save a small amount in the test tube.
15. Add three drops of vinegar to the mixture.
16. Record observations.
17. Add three drops of iodine solution to the mixture.
18. Record observations.
19. Place a small amount in the test tube and heat gently.
20. Record observations.

Observations and Data:

To Be Tested	Vinegar Fizzes	Iodine turns Blue	Litmus Test	Compound Melts
Cornstarch				
Sugar				
Baking Soda				
Mystery Mix				

Analysis and Questions:

1. How were you able to tell if all three compounds were not in your mystery mixture sample?

2. What would you conclude if you tested baking powder from your kitchen and found that it fizzed with vinegar, turned blue with iodine, but did not melt when heated?

Conclusion: Write a conclusion…Be sure to include the following: **objective, restate hypothesis**, **was your hypothesis correct**, **discuss results**, **and at least one thing you learned**.

APPENDIX C: Post-Survey

1. Enter Student ID#: ____________________

2. This part of the survey contains statements that relate to your feelings during this week's activity. There is no “right” or “wrong” answer. For each statement, think about the activity you just finished. Then, select the column that represents how you feel about each statement.

Statement	**I strongly agree**	**I agree**	**I don’t know**	**I disagree**	**I strongly disagree**
I was challenged and I felt I could meet the challenge.					
I did things naturally without thinking too much.					
I had a strong sense of what I wanted to do.					
I felt I was on track towards my goals.					
I was totally focused on what I was doing.					
I felt in control of what I was doing.					
It felt like nothing else mattered.					
I lost my normal sense of time.					
I really enjoyed what I was doing.					
I was in the zone.					

<u>Activity feedback</u>

3. At the end of the game, did you identify the correct thief on the first try?**   Yes [ ] No [ ]
4. Did this activity make you want to learn more about using science to solve a crime?   Yes [ ] No [ ]
5. What role did you play?**   Techie [ ] Pyro-technician [ ] Photographer [ ] Social networker [ ]

** question only included for students from the experimental group.

<u>Demographics</u>

6. Who is your science teacher?   Teacher A [ ] Teacher B [ ]
7. What period do you have science?   Period 1 [ ] Period 2 [ ] Period 5 [ ] Period 7 [ ] Period 8 [ ]
8. What is your gender?   Male [ ] Female [ ]
9. For this activity, was your group all girls, all boys, or a mix of both?   My group was all girls. [ ] My group was all boys. [ ] My group was a mix of both. [ ]

APPENDIX D: Revised Instruction Sheet for Mystery Powder Lab Activity

MYSTERY POWDER LAB ACTIVITY– INSTRUCTION SHEET

Objective: to determine the compounds in an unknown mixture by using the scientific method.

Materials:	test tubes (3)	test tube holder	test tube rack
	petri dish	eyedroppers (2)	vinegar
	iodine solution	water	pH paper
Powders:	sugar	cornstarch	baking soda

To begin, review your instructional sheet and complete DESCRIPTION OF EXPERIMENT and PLAN OUT YOUR EXPERIMENT.

Procedure for KNOWN POWDERS:

1. Place a scoopful (**One at a time)** of cornstarch, baking soda, and sugar in a **Petri Dish**. You must leave some of the mixture in the test tube for heating
2. Add three (3) drops of vinegar to each substance.
3. Record observations.
4. Wash and dry the **Petri dish**.
5. Place a scoopful of cornstarch, baking soda, and sugar in the **Petri Dish**.
6. Add three (3) drops of iodine solution to each substance.
7. Record observations in your DATA TABLE.
8. Wash and dry the **Petri Dish**.
9. Place a scoopful of cornstarch, baking soda, and sugar in the **Petri Dish**.
10. Add three drops of water to each substance. Complete pH paper test. Record your results in your DATA TABLE.

11. Take the remaining amount of each substance in the test tube. (Cornstarch in one, sugar in another etc.)
12. Gently heat the bottom of each test tube. Do not put the test tube directly in the flame.
13. Record observations in your DATA TABLE.

Review your observations and complete the
DATA TABLE SUMMARY and SYNTHESIS QUESTION.

Materials:	test tubes (1)	test tube holder	test tube rack
	petri dish	eyedroppers (2)	vinegar
	iodine solution	water	pH paper

Powder: mystery powder

Procedure for MYSTERY POWDER:

14. Pour mystery powder into the ***Petri Dish***. Be sure to save a small amount in the test tube.
15. Add three drops of vinegar to the mixture.
16. Record observations in your DATA TABLE.
17. Add three drops of iodine solution to the mixture.
18. Record observations in your DATA TABLE.
19. Add three drops of water to the mixture. Complete pH paper test.
20. Record observations in your DATA TABLE.
21. Place a small amount in the test tube and heat gently.
22. Record observations in your DATA TABLE.

To finish, review all of your observations and complete your
DATA SUMMARY (MYSTERY POWDER) and FINAL CONCLUSION.

APPENDIX E: Revised Investigation Sheet for Mystery Powder Lab Activity

MYSTERY POWDER LAB ACTIVITY	
Date	
Class Period	
Teacher	
Your Name	
Your Student ID#	
Group Members	

DESCRIPTION OF EXPERIMENT
In this space, describe the experiment that your group is going to perform. Explain as many details as you can.

PLAN OUT YOUR EXPERIMENT
In this space, write the steps that your group will perform to identify the mystery powder.

DATA TABLE (KNOWN POWDERS)				
	Vinegar Test	*Iodine Test*	*pH Test*	*Heat Test*
Cornstarch				
Sugar				
Baking Soda				

DATA TABLE SUMMARY: *For each test, describe any changes you observed with the powders.*

vinegar test:

iodine test:

pH test:

heat test:

SYNTHESIS QUESTION: *What will you be looking for when you test the mystery powder? Write a hypothesis for each test.*

If I add vinegar to the mystery powder then…

If I add iodine to the mystery powder then…

If I test the mystery powder with pH paper then…

If I heat the mystery powder then…

DATA TABLE (MYSTERY POWDER)				
	Vinegar Test	*Iodine Test*	*pH Test*	*Heat Test*
Mystery Powder				

DATA SUMMARY (MYSTERY POWDER)
Explain what each test revealed –if anything –about the mystery powder.

FINAL CONCLUSION
In this space, make your case for which compounds are included and not included in the mystery powder. *Be sure to state your conclusions based on the observations and data you collected.*

APPENDIX F: Incident Report

Northampton County / Incident Report Case Number: 2013-069216

EVENT INFORMATION	
Date	
Time Arrived	**AM / PM**
Location	
Lead Investigator	**Student ID#**
Co-Investigators	

DESCRIPTION OF INCIDENT
In this space, describe the incident your group is investigating. Explain as many details as you can. --- *Do you have any suspects yet? If so, support your claim with evidence.*

PLAN OUT YOUR INVESTIGATION
In this space, write the steps that your group will perform to identify the mystery powder.

EVIDENCE TABLE (KNOWN POWDERS)				
	Vinegar Test	*Iodine Test*	*Heat Test*	*pH Test*
Cornstarch collected at… ------------------------ *Location*				
Sugar collected at… ------------------- *Location*				
Baking Soda collected at… ------------------- *Location*				

EVIDENCE SUMMARY (KNOWN POWDERS): *For each test, describe any changes you observed with the powders.*

vinegar test:

iodine test:

heat test:

pH test:

SYNTHESIS QUESTION: *What will you be looking for when you test the mystery powder? Write a hypothesis for each test.*

If I add vinegar to the mystery powder then…

If I add iodine to the mystery powder then…

If I heat the mystery powder then…

If I test the mystery powder with pH paper then…

EVIDENCE TABLE (MYSTERY POWDER)				
	Vinegar Test	*Iodine Test*	*Heat Test*	*pH Test*
Mystery Powder *(obtained from detective)*				

EVIDENCE SUMMARY (MYSTERY POWDER)

Explain what each test revealed –if anything –about the mystery powder.

FINAL CONCLUSION

In this space, make your case for which compounds are included and not included in the mystery powder.
Be sure to support your conclusions based on the observations and evidence you collected.

Based on the evidence in your investigation, who is still a possible suspect? Circle all that apply.

Janitor *Secretary* *Destiny*

APPENDIX G: Support Letter

Bethlehem Area School District
Shaping the future. One child at a time.

Nitschmann Middle School

909 West Union Boulevard, Bethlehem, PA 18018
610-866-5781 • Fax 610-866-1435

April 23, 2013

Denise M. Bressler, Ph.D. Candidate
Teaching, Learning, and Technology Program
Lehigh University
E-104B Iacocca Hall, 111 Research Drive
Bethlehem, PA 18015

Dear Ms. Bressler:

On behalf of Nitschmann Middle School, I am very pleased to offer my commitment to your proposed dissertation project that focuses on enhancing science education by engaging middle school students through the use of a mobile learning game. Your advisor, Dr. Bodzin, has been an educational partner with our district for many years and his work has assisted us to promote inquiry-based learning, important spatial thinking skills, and science and environmental education with our students.

As part of your proposed project ***School Scene Investigators: The Case of the Mystery Powder,*** Nitschmann students will participate in the experimental intervention that will occur over a five-day period at the end of September. Specifically, classes of students will either be playing through the game or participating in the control curriculum. All students will fill out a short online survey and then write an incident report. A small sub-set of students will also be audiotaped. Students playing the game will use iPads supplied by Nitschmann Middle School.

The implementation will include all eighth grade science students with representation from different ability tracks. Nitschmann anticipates participation from approximately 300 diverse urban students. Nitschmann teachers will assist in this project by sending home informed consent forms and collecting them from students.

We are enthusiastic about the potential to partner with you on this project, and we look forward to participating in a meaningful learning experience that will successfully enrich our students' learning.

Sincerely,

Jacqueline Santanasto

Jacqueline Santanasto, Principal

APPENDIX H: Student Assent Form

LEHIGH
UNIVERSITY.

YES!

I'm willing to participate in the research study during the upcoming mystery powder lab activity.

Name: ____
(please print)

I know that I will be asked to fill out a short survey after the activity concludes and that the researcher will review copies of my incident report. I know that my identity will never be associated with the information that I provide. I also know that the researcher may observe my behavior in science class, take photographs, and audio record some of my conversations during class.

> NOTE: Photographs and audio tapes will be deleted after the research has been published in a peer-reviewed journal.

I have received permission from my parents and my consent form is attached.

____
Student Signature

____
Date

APPENDIX I: Informed Consent Form

Dear Parents or Guardians,

My name is Denise Bressler and I'm a doctoral student in Lehigh's College of Education. I am working with Dr. Alec Bodzin, professor of science and environmental education, on students' learning science using mobile technologies such as iPads. I will be conducting a small research study at your child's school as partial fulfillment of the requirements for my PhD degree in Learning Sciences and Technology. Students will be randomly assignment to participate in a control group with the usual science unit or an experimental group that will play a collaborative science mystery game using iPads supplied by the school.

This form is to request your permission to collect your child's responses to a few items that will measure learning outcomes. First, there will be a short survey. Your child's response to these surveys will not count towards his or her grade. Second, there will a lap report. Your child's written report will count towards his or her grade. Last, I will audio record some of the classroom conversations during the curriculum unit. Also, I request permission to observe your student during the curriculum and take some photographs. *NOTE: Some photographs may be taken for documentation purposes and for use in research articles. Any photographs will remain confidential with regard to your child's identity. Student names will not be associated with the photographs nor will the school name. Once collected, I will analyze all my data and write research articles about my findings. Photographs will be destroyed after articles are published in peer-reviewed journals.*

By granting me permission to examine your child's responses and learning behaviors, your child will be helping me to understand whether we can get students to learn more about science by playing science-based games with iPads.

Your decision about your child's participation in this study is voluntary. If you have any questions about this study, you may call me at 973-464-XXXX or email me at dmb309@lehigh.edu.

You may direct questions in regard to your child's rights pertaining to the use of the data in this study to the Office of Research and Sponsored Programs, Lehigh University, 610-758-3021. All reports or correspondence will be kept confidential.

To confirm your consent of your child's participation in this study, please sign below. Thank you.

Student name: ___________________________________

______________ ____

Date Signature of minor participant's parent or guardian

______________ ____

Date Investigator's signature

APPENDIX J: Grading Rubric for Culminating Artifact

GRADING RUBRIC FOR CULMINATING ARTIFACT STUDENT ID: ____

	Insufficient Response **0 points**	***Beginning*** **1 Point**	***Developing*** **2 Points**	***Proficient*** **3 Points**	***Exemplary*** **4 Points**	**PTS**
DESCRIPTION (P1: DEFINING THE PROBLEM)						
Stating the problem (P1a)	Student did not address the problem.	Statement is ambiguous about what is known and includes misunderstandings.	Statement vaguely establishes what is known.	Statement clearly establishes what is known but does not specifically discuss what is needed to be identified.	Statement clearly establishes what is known and specifically discusses or implies what is needed to be identified.	
				Experiment: Money is missing. There is a mystery powder. Identified correct suspects but no evidence to support. May have included their teacher as a suspect.		
PLAN (P3: PLAN AND CARRY OUT INVESTIGATION)						
Identify sequence of steps (P3)	Did not include a sequence of steps.	Formulates ambiguous steps that will be performed or procedure is incomplete (missing 3 or more steps).	Formulates general steps that will be performed, but missing 2 important steps.	Formulates general steps that will be performed but lacks sufficient detail.	Formulates extremely detailed steps that will be performed.	
KNOWN POWDERS (P4: ANALYZE/INTERPRET DATA)						
Produce data (P4a)	Did not complete the data table.	Includes results from only one test.	Includes results from 2 tests.	Includes results from 3-4 tests with basic information. May include some incorrect observations.	Includes results from 3-4 tests with detailed information. All observations are correct.	

KNOWN POWDERS - DATA/EVIDENCE SUMMARY (P4: ANALYZE/INTERPRET DATA)

Organize and summarize data (P4b)	Did not complete the data summary.	Described changes for 1-2 tests.	Described changes for at least 3 tests. May not be accurate.	Briefly described changes for 3-4 tests. Responses are not thorough. 1-2 test summaries are not accurate.	Thoroughly and accurately described changes for 3-4 tests as represented in their data table.

SYNTHESIS QUESTION (P1: ASKING QUESTIONS)

Asking questions (P1b)	Did not write any hypotheses.	Wrote hypothesis for 1-2 tests. Statements are missing important identifying components.	Wrote hypothesis for at least 3 tests. All statements are missing important identifying components.	Wrote a basic hypothesis for 3-4 tests. Some statements are missing important identifying components.	Wrote detailed hypothesis for 3-4 tests. Statements have important identifying components.

MYSTERY POWDER - DATA/EVIDENCE TABLE (P4: ANALYZE/INTERPRET DATA)

Produce data (P4a)	Did not complete the data table.	Includes results from only one test.	Includes results from 2 tests.	Includes results from 3-4 tests with very brief, basic information. May include some incorrect observations.	Includes results from 3-4 tests with detailed information. All observations are accurate.

MYSTERY POWDER - DATA/EVIDENCE SUMMARY (P6: CONSTRUCTING EXPLANATIONS)

Explanation of phenomena tested (P6)	No evidence summary included.	Summarizes the evidence table but does not articulate relevance of tests.	Explains the results of 1-2 tests and their general relevance.	Briefly explains the results of 3 or 4 tests and their relevance.	Thoroughly explains the results of 3 or 4 tests and their relevance.

FINAL CONCLUSION (P7: ENGAGING IN ARGUMENT FROM EVIDENCE)

Strength of argument (P7)	No conclusion stated or no correctly identified powder.	Identified only 1 correct powder.	Identified 2 correct powders but did not support with sufficient evidence. Third powder not mentioned.	Identified 2 correct powders supported with evidence. Third powder not mentioned or supported with evidence.	Identified 2 correct powders supported with evidence and used evidence to support why the other powder was not present.

OBTAINING, EVALUATING, AND COMMUNICATION INFORMATION (P8)

Evaluating and communicating scientific information (P8)	Student did not obtain, evaluate, or communicate information satisfactorily.	Identifies the claim(s). Has little or no evidence of purposeful organization.	Introduces the claim(s); organizes reasons and evidence in a manner that may lack cohesion (ideas may be rambling and/or repetitive); inconsistently uses words, phrases, and/or clauses to create cohesion and clarify the relationships among claim(s), reasons, and evidence.	Introduces the claim(s); organizes the reasons and evidence logically; uses words, phrases, and/or clauses to create cohesion and clarify the relationships among claim(s), reasons, and evidence.	Effectively introduces the claim(s); organizes the reasons and evidence logically in a manner that supports the writing task; effectively uses words, phrases, and/or clauses to create cohesion and clarify the relationships among claim(s), reasons, and evidence.

lacks cohesion in writing.

TOTAL POINTS:

APPENDIX K: Sample Student Responses for Culminating Artifact

Table A1

Student Response Examples from Both Groups for All Levels of Grading Rubric

Number	Rating	Quasi-Experimental Group	Control Group
Description of Incident/ Experiment (Practice 1)	4 pts	Theft at the cafeteria. White powder was left behind. Mr. Mayes wants our help with the investigation. Suspects: Destiny – working on a project with baking soda. Janitor – needs money for operation. Secretary – uses sugar when baking.	We are going to test each of the three substances in four different ways to identify the reaction in each one. After we do that, we will test the mystery powder to determine which two substances form it.
	3 pts	The cafeteria was robbed and there was mystery powder left behind. Suspects: Secretary because she baked cookies and needed sugar. Destiny because she needs baking soda for her science project.	There are mystery powders. Experimenting on known powders. We will use 4 tests to determine what the mystery powders are made of.
	2 pts	Someone stole money out of the cash register in the cafeteria. Suspects: Janitor needed money for surgery, the girl – she was baking cookies in the office cafeteria.	Our group is trying to find out what the mystery powder is made of.
	1 pt	Destiny got baking powder from the kitchen. Suspects: Destiny – she was the one who spilled the powder. Janitor – short on money.	Powdered sugar, baking soda, cornstarch, 4 tests.
Plan Out The Investigation (Practice 3)	4 pts	Check for the pH of the powders. Check for starch with iodine. Check if it's a base with vinegar. Check if it's flammable. Check which powder matches mystery powder.	First, we will test one of the known powders on all 4 tests. This process goes on until all of the 3 powders have been tested on all of the 4 tests. Second, we will test the mystery powder on all of the 4 tests. Third, we will use our observations to figure out what two powders mixed together make the reactions of the mystery powder.
	3 pts	Collect each powder. Identify mystery powder. Compare mystery powder and other powders. Accuse person.	Each person will take turns doing each test. Ex. I will do iodine, Dylan will do vinegar, etc. While one person is testing, the rest of the group will discuss and observe. When someone is done testing, we will quickly summarize and record data.
	2 pts	Question all the suspects. Test the powder.	We do all the tests on individual powders then test the mystery powder, then combos to find it.
	1 pt	Scan the codes of each person then talk to the person to find out what happened. Get some clues of the mystery to figure out who did it.	Each person in our group will do one of the four tests.

Evidence Table – Known Powders (Practice 4)	4 pts	All observations clear and accurate.	
	3 pts	Most observations clear and accurate. Generally missing pH numbers - only providing a color.	
	2 pts	Incomplete data table with errors. Ex. Cornstarch vinegar test: It started to turn red. Generally missing pH numbers, only reported a color.	
	1 pt	Mostly missing or unclear observations.	
Evidence Summary – Known Powders (Practice 4)	4 pts	Vinegar: cornstarch and sugar turned into a thick liquid and baking soda fizzed up. Iodine: cornstarch and sugar turned into a liquid but cornstarch changed color. Baking soda turned into orange paste. Heat: Baking soda did not change. Sugar turned into orange syrup. Cornstarch burned. pH: Cornstarch is more acidic while sugar is neutral and baking soda is basic.	Vinegar: with baking soda it bubbled and foamed. Nothing really changed the cornstarch and sugar. Iodine: with the cornstarch it absorbed and made it stick together. It made sugar and baking soda clump and mush. Heat: Cornstarch and sugar both burned and changed to black or brown. The baking soda smoked. pH: baking soda was the lightest color (yellow). Cornstarch and sugar were darker colors.
	3 pts	Vinegar: cornstarch and baking soda fizzed. Iodine: cornstarch was purple, baking soda was orange, sugar was nothing. Heat: cornstarch burns and sugar melts. pH: cornstarch and sugar was 6 and baking soda was 8.	Vinegar: during the vinegar test two of them stayed the same and one started to foam. Iodine: they all turned a color. One was hard solid, one was a jelly, another was hard. pH: they all turned a different color. Heat: they started to melt and become a liquid except baking soda.
	2 pts	Vinegar: the baking soda bubbles, when we put the vinegar. Iodine: the iodine hardened, slides off, or turned to slime. Heat: the sugar began to burn and smoke	Vinegar: baking soda bubbled. Iodine: cornstarch turned into tar-ish substance. pH: all were 6, 7, or 8. Heat: cornstarch smelled like marshmallow.
	1 pt	Vinegar: The vinegar to rise and the texture started to change. Iodine: the baking soda changed colors, turned orange. Heat: it looked like it is started to burn sugar. pH: it is starting to rise.	Vinegar: The cornstarch turned into hard balls and sugar turned like a glue. Baking soda fizzed up when we put vinegar on it.
Synthesis Question (Practice 1)	4 pts	If I add __ to the mystery powder then… Vinegar: cornstarch would not fizz, sugar would not fizz, baking soda will fizz. Iodine: sugar will turn clear, cornstarch will turn dark blue, baking soda turns orange. Heat: cornstarch will turn black, sugar will melt, baking soda will stay the same pH paper: baking soda will show up yellow (ph8), cornstarch and sugar will turn orange (ph6)	If I add __ to the mystery powder then… Vinegar: it will bubble if there is baking soda. Iodine: it will turn crumbly and black if there is cornstarch in it. Heat: it will bubble, expand, and smoke if there is sugar pH paper: it is cornstarch if you get ph 6
	3 pts	If I add __ to the mystery powder then… Vinegar: it will bubble if there is baking soda in it. Iodine: it will crumble if there is cornstarch in it. Heat: It will burn if there is cornstarch or sugar in it. pH paper: it will react to pH6.	If I add __ to the mystery powder then… Vinegar: it will bubble if it contains baking soda Iodine: it will clump without sugar Heat: it will do something without baking soda pH paper: it will be high without sugar

	2 pts	If I add __ to the mystery powder then… Vinegar: it will fizz. Iodine: it will turn purple Heat: it will burn pH paper: 6	If I add __ to the mystery powder then… Vinegar: it will bubble and expand. Iodine: it will stain it and nothing will happen Heat: it will turn black, bubble and burn up pH paper: it will turn to mixed colors
	1 pt	If I add __ to the mystery powder then… Iodine: I would be looking for it to turn dark purple or black. Heat: I would be looking for changes with the brown liquid.	If I add __ to the mystery powder then… Vinegar: it would bubble and fizz because of the baking soda.
Evidence Table – Mystery Powder (Practice 4)	4 pts	[all observations accurate and clear] Vinegar: it fizzed Iodine: turned black Heat: turned black at bottom pH paper: pH 8	[all observations accurate and clear] Vinegar: it bubbled and mixed together Iodine: it repelled & turned black like cornstarch Heat: it burned on the bottom; it turned black, yellow, brown. It rose. It separated towards the top. It started smoking. pH paper: pH 9 – green – like baking soda
	3 pts	[1 unclear observation] Vinegar: bubbling, fizz, foam Iodine: became clear *Heat: took a long time to change* pH paper: became pH 8	[1 unclear observation] Vinegar: It started to bubble Iodine: turned into a solid and turned black. Heat: it's smoking, turning black, rose up. *pH paper: turned into a solid*
	2 pts	[2 unclear observations] Vinegar: made the powder fizz up *Iodine: it changed colors* Heat: it was starting to burn the powder *pH paper: it turned yellow. It's watery.*	[2 unclear observations] Vinegar: It started to bubble Iodine: It got hard and turned black. *Heat:* *pH paper: It got hard.*
	1 pt	[Not applicable. This score was never assigned.]	[3 unclear observations] *Vinegar: bounced off* *Iodine: sticked.* Heat: smoky, smells like bonfire *pH paper: looks like dough.*
Evidence Summary – Mystery Powder (Practice 6)	4 pts	The vinegar test made it fizz. It matched cornstarch and baking soda. It turned dark purple and it matched cornstarch. It burned and match cornstarch. But the pH was 7 and it matched sugar.	From doing the vinegar test, the mystery powder bubbled like the baking soda. After doing the iodine test, the mystery powder turned gooey like the sugar. For both the mystery powder and the sugar, the pH test was pH7. After the heat test, the mystery powder smoked, bubbled, and turned black like the sugar.
	3 pts	Vinegar and heat matched cornstarch. pH matches sugar. Iodine matches baking soda.	When adding water to it, it turned flubbery like cornstarch did. When adding vinegar to the mystery powder it fizzed like baking soda did. Iodine test made it turn into a black tarry substance.
	2 pts	Baking soda fizzes when vinegar hit it. Cornstarch burned. The pH test turned green.	For the vinegar test, the mystery powder began to bubble and then the vinegar rolled off of the powder. In the iodine test the powder turn brown and hardened. The pH level is 8, the same as baking soda. In the heat test, the powder turned black. It had a weird smell to it.

	1 pt	The tests revealed that the powder fizzes, turns orange, burns, and pH test of 7.	Vinegar bubbled and looked like glue. Iodine sinks through. Heat test moves the powder up.
Final Conclusion (Practice 7)	4 pts	In the heat test, it turns black at the bottem – not sugar. In the vinegar test, the powder started to fizz witch means that it contains baking soda. And in the iodine test, the powder turned black, so it is a mixture of baking soda and cornstarch.	There was not any sugar in it because if there was the iodine and the vinegar would've dissolved in it. The reactions the mystery powder showed were exactly the same to cornstarch and baking soda, by bubbling up and the iodine rolling over.
	3 pts	The mystery powder is a mix of cornstarch and baking soda. It fizzed which only baking soda does but it burned which only cornstarch does.	Our research showed that the traits of the mystery powder match those of cornstarch and baking soda. The vinegar test fizzes and bubbles like baking soda. Also, with the iodine, it turned into a tarish black substance similar to the reaction of cornstarch.
	2 pts	I think baking soda and cornstarch.	The mystery powder is baking soda and cornstarch because of the reactions that occurred.
	1 pt	The mystery powder might be sugar.	I think the mystery powder is cornstarch.

APPENDIX L: A Priori Codes for Qualitative Data Analysis

In order to analyze the conversational dialogue, two different sets of codes were used. The first chart details codes that supported the investigation of the 3rd research question: How do in-team communication responses during the INPLACE mobile game compare to in-team responses during the control activity? Building on the work of Barron (2003), I utilized her code structure of accept, discuss, and reject.

Peer Responses	Codes with definitions
Accept	PR_Accept: Possible student interactions include agreeing with the speaker, supporting the proposal, proposing a next step, documenting the proposal, or assisting with documenting the proposal.
Discuss	PR_Discuss: Possible student interactions include any interaction that would facilitate further discussion. This could include questioning a proposal, challenging a proposal with new information, or requesting clarification.
Reject	PR_Reject: Possible student interactions include any interaction that would not facilitate further discussion. This could include providing no verbal response, ignoring a proposal, and rejecting a proposal.

The second chart details codes that supported the investigation of the 4th research question: How do scientific practices of teams in the INPLACE mobile game compare to those of teams in the control activity? Squire and Jan (2007) determined that mobile AR games can promote scientific argumentation; their codes included question, hypothesis, counter-hypothesis, and evidence. I expanded on this code structure and utilized codes that aligned directly to the Next Generation Science Standards.

Scientific Practice	Codes with definitions
Practice 1: Asking Questions/ Defining Problems	SP_1: Student dialogue includes asking a question pertaining to the scientific investigation, establishing what is already known about the scientific investigation, or determining what still needs to be answered.
Practice 3: Planning Out Investigation	SP_3: Student dialogue discusses the investigation plan or data collection. Students may also identify what is to be recorded or how it will be measured.
Practice 4: Analyzing and Interpreting Data	SP_4: Student dialogue interprets relationships between data. This could include discussing characteristics of the data collected or similarities and differences in the data.
Practice 6: Constructing Explanations	SP_6: Student dialogue explains relationships between data with qualitative or quantitative reasons. Ideally, students are constructing a scientific explanation based on valid and reliable evidence obtained from sources.
Practice 7: Engaging in Argument from Evidence	SP_7: Student dialogue includes convincing arguments that support or refute explanations. Student dialogue 1) clearly cites relevant evidence or 2) poses or responds to questions with evidence.

Appendix M: Definition of Terms

Augmented reality (AR). Augmented reality is when the physical environment is enhanced with digital information. Specifically, in this study, iPads enabled the students to experience augmented reality, or the digital information that was embedded in their school environment.

Command language. As defined by this study's author, command language is a type of directive language used mostly by control teams. Examples included "put the whole entire thing in" and "go get the other one."

Communal language. As defined by this study's author, communal language is a language style in which members address the team collectively and refer to the team as an entity with words such as "we," "we're," and "let's." Game teams used this language style more frequently than control teams.

Engaged responses. According to Barron (2003), when students accept or discuss a teammate's proposal, then they are offering engaged responses.

Engagement. Student engagement occurs when students are attentive, curious, interested, and possibly passionate about their learning. Students may show high levels of motivation and optimism as well.

Epistemic game. According to Shaffer (2006), epistemic games are games in which players inhabit professional roles and use the toolkit of that profession to solve authentic problems.

Flow. Flow is a positive psychological state that is challenging, intrinsically rewarding, and enjoyable (Csikszentmihalyi, 1975). Csikszentmihalyi has posited that there are nine elements of the flow experience which include the right challenge-skill balance, sense of control, and enjoyment.

Game-based learning. Simply put, games used for educational purposes could be categorized as game-based learning; this is one definition used within this study. However, the term game-based learning is also used to describe the effective learning experienced during gameplay.

Gamefulness. The term gamefulness has arisen in contrast to gamification. While gamification is about points, leveling up, and badges…gamefulness is about what is intrinsically appealing and inspiring about the content.

Hard fun. According to Papert (2002), everyone enjoys hard, challenging things to do provided the challenge is well-matched to the individual; Papert coined the phrase "hard fun" to encapsulate the notion that desirable activities are both challenging and enjoyable.

INPLACE. This is an acronym invented by this study's author that stands for **I**nterdependent, **N**etworked, **P**articipatory **L**earning, **A**ugmented, **C**ollaborative **E**xperience; it is used to describe the type of mobile game created for this study.

Non-engaged responses. According to Barron (2003), when students reject another teammate's proposal, they are offering non-engaged responses.

Scientific practices. According to the 2012 report published by the National Research Council, science education should be designed around eight scientific practices as outlined in the report. The Next Generation Science Standards (NGSS) has created educational standards designed to support these scientific practices which include more emphasis on higher-order thinking skills, rather than fact memorization.

Serious Educational Games (SEGs). Since 2003, there has been a movement to use video game technology for learning. The commercial games and grassroots games that are used for learning and that are well-designed games (rather than interactive worksheets) are called Serious Educational Games, or SEGs for short.

21st century learning skills. This set of skills includes critical thinking, problem solving, communication, and collaboration. Such skills and expertise are essential for today's work-life.

DENISE MARIE BRESSLER

Biography

Throughout her research, writing and professional career, Denise has always been interested in the affordances of innovative technologies and how they can best be used as effective educational tools. In 1997, for her undergraduate thesis at Princeton University, she conducted a meta-analysis of internet usage in K-12 education and determined that several variables showed positive correlations with improved student learning: interactivity with analysis, the subject of science and working in groups. In 2000, as a master's student at Teachers College, Columbia University, she built a CD-ROM game grounded in educational theory because she saw the potential of technology to personalize learning through intelligent-computer assisted instruction.

In 2000, as Denise began her career, dot-coms were on the rise and her first job was with MaMaMedia.com, an educational website for children based on constructionist learning. Ultimately, she found a happy home for over 6 years at Liberty Science Center in Jersey City, NJ. There, she studied the current technology trends of the time and conceived of a mobile learning project because she hoped that if visitors used their mobile phones in her exhibits it would inspire them to continue learning once they left the museum. Finally, as a doctoral student in Lehigh University's College of Education, she has seen that schools are beginning to embrace mobile devices but there are so few pedagogical models to follow, so she built a series of mobile science games called *School Scene Investigators* to engage learners through feelings of flow and successful collaboration. As Dr. Denise Bressler moves forward in her career, she intends to keep creating innovative learning experiences based on what the next new tool can best provide.

EDUCATION

Lehigh University	Ph.D. in Learning Sciences and Technology	2014
Teachers College Columbia University	M.A. in Instructional Technology and Media	2000
Princeton University	A.B. in Sociology	1997

RECENT PUBLICATIONS

Bressler, D. M. (in press). School Scene Investigators: A collaborative AR game for middle school science inquiry. In C. Holden, S. Dikkers, J. Martin, & B. Litts (Eds.), *Mobile Media Learning: Innovations and Iterations*.

Bressler, D., & Bodzin, A. (2013). A mixed methods assessment of students' flow experience during a mobile augmented reality science game. *Journal of Computer Assisted Learning, 29*(6), 505-517.

Bressler, D. M. (2013). Museums: Gateways to mobile learning. In Z. Berge & L. Muilenburg (Eds.), *Handbook of Mobile Learning* (pp. 224-233). New York: Routledge.