Exploring the Impact of Varying Levels of Augmented Reality

to Teach Probability and Sampling with a Mobile Device

by

Quincy Conley

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved October 2013 by the
Graduate Supervisory Committee:

Robert K. Atkinson, Chair
Frank Nguyen
Brian C. Nelson

ARIZONA STATE UNIVERSITY

December 2013

ABSTRACT

Statistics is taught at every level of education, yet teachers often have to assume their

students have no knowledge of statistics and start from scratch each time they set out to

teach statistics. The motivation for this experimental study comes from interest in

exploring educational applications of augmented reality (AR) delivered via mobile

technology that could potentially provide rich, contextualized learning for understanding

concepts related to statistics education. This study examined the effects of AR

experiences for learning basic statistical concepts. Using a 3 x 2 research design, this

study compared learning gains of 252 undergraduate and graduate students from a pre-

and posttest given before and after interacting with one of three types of augmented

reality experiences, a high AR experience (interacting with three dimensional images

coupled with movement through a physical space), a low AR experience (interacting with

three dimensional images without movement), or no AR experience (two dimensional

images without movement). Two levels of collaboration (pairs and no pairs) were also

included. Additionally, student perceptions toward collaboration opportunities and

engagement were compared across the six treatment conditions. Other demographic

information collected included the students' previous statistics experience, as well as

their comfort level in using mobile devices. The moderating variables included prior

knowledge (high, average, and low) as measured by the student's pretest score. Taking

into account prior knowledge, students with low prior knowledge assigned to either high

or low AR experience had statistically significant higher learning gains than those

assigned to a no AR experience. On the other hand, the results showed no statistical

significance between students assigned to work individually versus in pairs. Students

i

assigned to both high and low AR experience perceived a statistically significant higher level of engagement than their no AR counterparts. Students with low prior knowledge benefited the most from the high AR condition in learning gains. Overall, the AR application did well for providing a hands-on experience working with statistical data. Further research on AR and its relationship to spatial cognition, situated learning, high order skill development, performance support, and other classroom applications for learning is still needed.

DEDICATION

More than to anyone else, I dedicate this accomplishment to Jodi, my life-long partner, for the constant love and support she provided me throughout this ordeal. Without her relentless (and I do mean relentless) encouragement, support, and understanding, undoubtedly, none of this would have been possible. No matter the difficulties I faced, she is the one who told me never to give up on obtaining "the blue vase." My blue vase is dedicated to her!

ACKNOWLEDGMENTS

To my core, I want to thank my mother who taught me "success is up to you." I also want to thank my brother who, even though he is younger than me, has always been my role model. To my grandmother and grandfather, thank you for your continuous words of encouragement, inspiration, and motivation. All the things that went into making this happen, I did this for you, my family.

A special thank you to my dissertation committee for shaping and contributing to the success of this adventure. My utmost gratitude is due to the chair of my committee, Dr. Robert K. Atkinson. Thank you for providing so many amazing experiences. I especially appreciated the wonderland of research tools that you provided to me. To Dr. Frank Nguyen, who has not only been supportive of my research pursuits but also my growth as an aspiring professional in academia. Thank you for taking an interest in me and believing in me. To Dr. Brian Nelson, thank you for your valuable suggestions and the advice that you have shared over this process.

I want to especially thank Ben Boren, Benton Redmann, David Magill, Ivan Montiel, and Ryan Burgoyne, who developed the main learning intervention used for this study. Thank you for building the best mobile app that an aspiring researcher could ever ask for.

Also, I will always remember the help from my classmates in and around the Educational Technology program at ASU. I will forever be indebted to the impact and value of my classmates' support. Having the opportunity to work with each and every one of you has been amazing. No matter if it was about classwork, concerns about my

research, or other parts of life, thank you for all of your help. I look forward to continuing our relationships.

Also, I would like to express my sincere gratitude to all my friends who encouraged and supported me along the way. You made a difference.

To my Kappa Delta Pi (Honor Society) crew, thank you for taking me in and allowing me to be a part of something so important. Being a member of KDP has taught me what real service is.

The following list is undoubtedly incomplete. I want to thank all of those who have gone out of their way to support me along my journey: Dr. Gary Bitter, Dr. Willi Savenye, Dr. Mark Grabe, my Michaels & Associates family, Dr. Ann Ewbank, Dr. Martha Cocchiarella, Dr. Janel White-Taylor, Dr. Jill Oliver, Dr. Sam Green, Dr. Jim Klein, Dr. Jill Stamm, Dr. Marylin Thompson, Dr. Leanna Archambault, Dawn Durante, Jason Bradshaw, Tim Duarte, Kia Velarde, Jorge Garcia-Ramirez, and many more.

As I move on to the next chapter of my life, may I not just stand on, but may I "crush the shoulders of giants" who have led the way before me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

The General Problem

Statistics is an important area of study, which impacts virtually every aspect of our daily lives. No matter if it is an elementary student who is just starting out learning fractions, or a postdoctorate who uses higher-order statistics to make sense of the world around us, the need to understand statistics is omnipresent (Batanero & Diaz, 2011). Students enrolled in a wide range of classes such as elementary school math, high school algebra, or college statistics are often asked to identify important characteristics of a group of items, calculate a proportion, or compare groups of different sizes to each other. In other words, basic statistic skills are necessary for students at all levels. For example, such skills are required to analyze a sample by estimating the probability of a certain characteristic of an item from a sample representative of a population or to examine the variability between samples. However, effectively learning statistical reasoning skills like these fundamental concepts remains elusive (Garfield & Ben-Zvi, 2007).

Drawing from the scholarly literature, it seems that across all levels of education, teachers frequently must assume that students start without knowledge of statistics; in essence, instructors start from scratch each time they set out to teach statistics (Batanero & Diaz, 2011; Garfield & Ben-Zvi, 2004; Hirsch & O'Donnell, 2001; Jones, Langrall, & Mooney, 2007). One of the primary challenges with teaching these essential statistical concepts is to help students understand how different statistical decisions and techniques affect the kinds of conclusions that may be drawn from it. Historically, statistical concepts have been taught in two parts: (1) students are given the theoretical perspective

1

in the form of a lecture or a reading assignment, and (2) they are given homework to practice applying the theory. The issue with this standard, albeit passive, method of instruction is that it requires the students to understand the theory separately from the context in which they will practice in most cases (Chance, 2002; Garfield, 1995, 2002; Lovett, 2001; Sedlmeier, 1999). This is particularly problematic when teachers are instructing students about the abstract concepts behind probability and sampling, which is often difficult to teach due to students' basic statistical reasoning skills have lacked development over the course of their schooling (Cobb & Moore, 1997; Garfield & Ben-Zvi, 2007; Snee, 1993). Moreover, teaching basic statistical concepts is difficult due to the lack of the hands-on application and practice of the basic mathematical computation techniques (Chance, 2002; Garfield, 1995, 2002; Lovett, 2001; Sedlmeier, 1999; Shaughnessy, 2007). In most cases, students need to be able to calculate a proportion in order to learn these important skills, which are considered fundamental to solving probability and sampling-type problems. Another complication with teaching statistics is students' inability to conduct mathematical calculations (Garfield, 2002; Kahneman, Slovic, & Tversky, 1982). In most cases, students are required to calculate proportions as percentages and decimals, which is a difficult task for novice mathematicians. Researchers recognize this problem and suggest simplifying the task by using counts and ratios rather than percentages and decimals to help students correctly develop their statistical reasoning skills (Garfield & Ben-Zvi, 2007).

 With regard to teaching statistical reasoning, most studies focus on topics related to probability (Chance, Ben-Zvi, Garfield, & Medina, 2007; Dubois, 2006; Garfield & Ahlgren, 1988; Hirsch & O'Donnell, 2001; Lovett, 2001; Moore, 1997; Shaughnessy,

2007); however, there is a growing focus on additional topics such as sampling, variability, and the role of technology (Chance, Delmas, & Garfield, 2004; Garfield & Ben-Zvi, 2007; Pratt, Davies, & Connor, 2011; Shafto & Goodman, 2008). Although some instructional strategies lead to positive results, researchers continue to seek better ways to teach statistical reasoning. Fueling this movement, researchers continue to emphasize the value of statistics and the importance of improving how it is taught across all levels of education (Shaughnessy, 2007). While students can learn how to compute formal measures of inferential statistics, they rarely understand what these summary statistics represent, either numerically or graphically, and do not understand their importance and connection to other statistical concepts. Despite the widespread belief that statistical reasoning is a critical skill, current research continues to struggle with ways to better help students to reason about probability, sampling, and variability (Garfield & Ben-Zvi, 2004). One solution is to consider new methodological frameworks to improve upon what statistics education experts and educational technologists currently understand about best practices to teach statistical reasoning. In particular, there seems to be an opportunity to explore ways where students with basic mathematical skills are able to engage in interactive practice opportunities that allow them to explore the fundamental theoretical concepts underlying statistical reasoning in the context of concrete, easy-to-understand examples.

Framework of the Study

This study proposes a new instructional framework for teaching statistical reasoning skills that leverages advancements in augmented reality (AR) and mobile technologies in combination with three main pedagogical approaches: collaborative

learning, embodied cognition, and situated learning. The framework is designed to encourage students to carry out a contextualized scientific inquiry with interactive practice opportunities for statistical theoretical concepts using the emerging technology of AR and mobile devices. AR was selected as one of the foundational technologies for this study due to its ability to "create an artificial world" by superimposing digital objects on the real world (Höllerer & Feiner, 2004, p. 1). Mobile technology was also selected in order to allow students to move freely about a digital space collaboratively while completing a learning activity, which was designed to resemble authentic statistical practice (Roseth, Garfield, & Ben-Zvi, 2008). The research literature in this area suggests that mobile-enabled AR experiences can uniquely provide authentic learning environments that are potentially more engaging than traditional educational settings (Huizenga, Admiraal, Akkerman, & Dam, 2009). Collaborative learning was selected for inclusion in the study's framework given that this approach can facilitate transfer and knowledge acquisition without the need for heavy oversight from an instructor (Dillenbourg, Järvelä, & Fischer, 2009; Gredler, 2011; Vygotsky, 1978). Based on embodied cognition, the framework includes three levels of AR experience: a high AR experience (interacting with three-dimensional, or 3-D images coupled with movement throughout a physical space), a low AR experience (interacting with 3-D images without movement throughout a physical space), and no AR experience (interacting with two-dimensional, or 2-D, images and text without movement throughout a physical space) (Billinghurst & Kato, 2002). Based on the body of literature from the situated learning theory domain, the learning activity was situated in a specific context with the desired goal of providing a more concrete definition for abstract concepts such as statistical

reasoning. Specifically, during the learning activity, the aim was for the student to perceive him or herself immersed in a real-world environment. In sum, this framework—an extension of the existing research literature—was used to validate whether AR and mobile technology can be used to cognitively anchor contextualized information through movement to enhance students' engagement levels and learning gains (Price & Rogers, 2004).

Challenges for Developing Statistical Reasoning Skills

As a major subset of statistics education, Gal and Garfield (1997) defined statistical reasoning as "the way people reason with statistical ideas and make sense of statistical information" (p. 207). In essence, statistical reasoning consists of making interpretations and statistical summaries based on sets of data. Statistical reasoning is a concept worth distinguishing from its counterpart, statistical processes (i.e., mean, mode, standard deviation, etc.). This distinction is consistent with the literature on statistics education, which separates statistical reasoning from statistical process, with the latter more focused on the computation behind statistical reasoning (Chance, 2002). Moreover, statistical reasoning is defined as the formulation of ideas about data, which leads to statistical inferences and results (Garfield, 2002). As a result, the focus of most of the recent research related to statistical reasoning covers topics such as statistical sampling, probability, proportions, and variability (Garfield & Ben-Zvi, 2007). Meletiou-Mavrotheris and Lee (2002) conducted a study evaluating whether students' statistical reasoning could be improved by requiring them to construct inferences and results from given data. During this study, undergraduate students' statistical reasoning skills were observed and analyzed over a portion of a semester in a university-level introductory

statistics course. Using a pretest to posttest comparison method, students' learning gains were evaluated on their understanding at the beginning and end of the instruction. The researchers found that students demonstrated an increased understanding on tasks that required statistical reasoning skills, such as making statistical inferences from a set of data drawn at random from a particular population. The researchers attributed the positive results to moving the statistical concept of variation from the periphery to a more central focus of statistics instruction in the form of real-world interpretive contexts. By pointing out sampling variability in a real situation that was considered relevant to the students, they were better able to recognize variations and make appropriate inferences about a sample—thus, performing better at the end of the introductory course (Meletiou-Mavrotheris & Lee, 2002).

Garfield, delMas, and Chance (2007) used a classroom-based research design to test and revise a lesson to help students develop reasoning about variability. In the study, a group of novice and experienced teachers collectively designed a lesson to help reveal and build on students' informal assumptions about variability. The authors found that the sequence of activities helped students develop a deep understanding of the concept of variability as well as understanding of measures such as range, mode, and median. Instead of assigning homework to be done outside of class after a typical lecture, students were first introduced to a statistical concept by reviewing the course material digitally using a CD-ROM. The initial review session provided the students with an overview, key definitions, and examples of the concept. Then, the valuable class time was spent with the instructor answering critical questions and filling in gaps in the students' understanding. During the class time, students were also given the opportunity to work in groups to

complete hands-on activities and computer lab exercises. The activities were customized using an iterative strategy with the purpose of increasing students' statistical reasoning skills by building on what they already knew about the content. Garfield et al. (2007) concluded that statistics education experts should consider new and innovative methods of instruction for teaching an introductory college course. Moreover, a more thoughtful approach should be considered for students who enter the course with low statistical reasoning skills. Garfield et al. (2007) make the case that the new framework to teach sampling, probability, and variability should deviate from the traditional linear progression most college statistics courses follow. Notwithstanding the efforts of notable studies in this realm, there is a lack of innovation in the way statistical education is currently taught in practice. "Change is never easy;" however, a considerable change in content delivery is needed, or statistics education risks being even further neglected by future generations (Snee, 1993, p. 151). Ultimately, lack of statistical knowledge and interest could adversely affect the number of students who pursue degrees in science, technology, engineering, and mathematics (STEM) fields where statistics education is an integral part of the curriculum. There is already evidence that this is happening in the STEM fields, (Watkins & Mazur, 2013). Unless this trend is reversed, it is plausible that future generations will simply avoid these fields because they are not adequately prepared with the proper prior knowledge.

Collaborative Learning in Statistics Education

How can students learn to reason statistically? In general, it is thought that students learn best when given the opportunity to struggle with their own understanding of an unfamiliar concept, which often happens by reducing reliance on lecturing as the

7

primary means of instruction and instead providing more structured collaborative

learning activities (Garfield & Ben-Zvi, 2007). In support of this notion, Franklin and

Garfield (2006) advocated to the National Council of Teachers of Mathematics that, "as a

rule, teachers of statistics should rely much less on lecturing, and much more on the

alternatives such as projects, lab exercises, and group problem solving and discussion

activities" (p. 345).

Correspondingly, statistics education experts have suggested that learning

outcomes can be directly beneficial when students actively work together with others

where they can share ideas, resolve conflicting beliefs, and solve problems to construct

their own understanding of the information (Garfield & Ben-Zvi, 2007; Lin, 2011;

Moore, 1997). Researchers have explored the impact and value of group work related to

teaching statistics in an effort to identify effective learning strategies (Garfield & Ben-

Zvi, 2008; Kaufmann & Schmalstieg, 2003; Moore, 1997; Roseth, Garfield, & Ben-Zvi,

2008; Singleton, 1989). These studies primarily evaluated groups working in class to

solve a problem, discuss a procedure, or analyze a set of data.

What does it mean to work cooperatively? The idea is broad, however, one

pedagogical theorist close to this domain area summed it up within collaborative learning

theory. Dillenbourg (1999) described collaborative learning theory as "the criteria of the

situation…, the interactions…, processes…, and…effects" for working toward a common

goal (p. 13). The goal can vary, but in the context of this study, collaborative learning is

for the purpose of gaining knowledge. According to Dillenbourg's definition, there is a

relationship between the four criteria. The situation dictates the division of

responsibilities and interactions patterns, and the interactions generate cognitive

occurrences, which can lead to learning (Dillenbourg, 1999). In short, it is important to note that the goal of collaborative learning theory should be to improve the learning of each individual student while working as a team. All tools and resources within the learning activity should be designed to produce collaboration that scaffolds and supports scientific thinking (Dillenbourg et al., 2009). Therefore, collaborative learning was one of the important guiding theories to examine the effects of cooperation on learning outcomes in this study.

Chick and Watson (2001) conducted multiple studies investigating the impact of collaborative learning for teaching statistics. In one study, they investigated different levels of cooperative learning in teaching statistics to grade school students, and they found generally positive results. Specifically, students working in small groups of three were observed solving statistical problems. Those who were observed as having positive collaboration experiences had positive views towards group work. This is an important consideration in the context of this study, because the design of collaborative instruction should result in students' increased learning gains given a cooperative learning experience. Similarly, delMas, Garfield, and Chance (1999) examined students' abilities to reason about sampling distributions while working in groups in an introductory university-level statistics course. They found that students' reasoning about sampling distributions improved as the activity was changed to embed guiding assessment questions within the activity and when the activity had students make statistical inferences about different sampling distributions from various populations. This study was later replicated in a different type of undergraduate course and similar results were found (Lunsford, Holmes-Rowell, & Goodson-Espy, 2006). Lunsford et al. investigated

9

their understanding of sampling and probability with undergraduate math students in two introductory level algebra-based statistics courses. The students' performance was assessed after receiving a half of a semester's instruction comprised of directed instruction with hands-on activities and simulations. Comparing the pre- to posttest measurements, students scored significantly higher on the second test overall. However, the researchers reported that even though the students demonstrated the ability to complete the necessary computations, they were not as proficient at answering questions that required statistical reasoning skills. Their conclusion was that students needed more practical application experience that included "graphical representations" of the sampling and probability concepts to develop their statistical thinking ability (Lunsford et al., 2006, p. 19). Research has indicated that collaborative learning—with or without technology— is a tempting phenomenon, but high-level collaboration among students in real-life learning settings (i.e., in classrooms) is more difficult to realize than previously thought (Häkkinen & Hämäläinen, 2012).

While there are many aspects to consider when designing an ideal collaborative learning experience, (i.e. the situation, number of interactions, roles and responsibilities, and types of communication), this study focuses on group size given the critical role it plays in the outcome of collaborative learning experiences. What size collaborative team produces the optimal effect in statistics education? There are surprisingly few studies that directly examine the ideal group size for collaborative learning in statistics education; however, there are a substantial number of experts that support integrating collaborative learning opportunities into how statistical reasoning is taught (Giraud, 1997; Keeler & Steinhorst, 1995; Lovett, 2001; Magel, 1998). These researchers postulated that

10

collaborative learning helps students learn statistics in ways that not only enhance their statistical reasoning and communication skills, but also gives them practice working collaboratively, which models the collaborative nature of real statistical work. Supported in an article by Roseth et al. (2008), cooperative learning in small groups can also enhance critical thinking, conceptual understanding, and other higher order skills.

The definition of *small* group size ranges from as few as two or as many as five students (Dillenbourg, 1999). Across most disciplines, educators realize that as the size of the learning group increases, the likelihood of the deep understanding decreases. Also, the larger the group size, the more instructional guidance required to help the group succeed (Roseth et al., 2008). Larger groups can reduce the opportunity for each student to contribute and collaborate, whereas in smaller groups it is difficult for students to hide from participating in the activity. It actually forces students to make more of a contribution and share their ideas just by the dynamics of working in small groups. From an educator's perspective, the smaller the group, the easier it is to identify difficulties and distractions. Therefore, smaller group sizes are preferred to improve the effects of the instruction (Kreijns, Kirschner, & Jochems, 2003). The consensus from the studies previously reviewed is that working in groups helps students become more involved in their own learning. However, group work should not be a part of teaching statistics just for the sake of cooperation. Instead, it should be thought of as an integral part of helping students learn. Without good instruction that is carefully designed with guidance and supervision, collaborative learning will not be effective (Garfield & Ben-Zvi, 2007).

There are also a few studies that consider group performance and behaviors in the classroom while using certain technologies such as mobile devices or games (Morgan &

Butler, 2009; Schwabe, Göth, & Frohberg, 2005; Wagner, Pintaric, Ledermann, &

Schmalstieg, 2005). Wagner et al. (2005) suggested that AR and mobile technologies

complement each other for presenting natural 3-D information while moving through a

physical space for collaborative, multi-user experiences. Morgan and Butler (2009)

posited that mobile devices with large displays were ideal for collaboration, because

learners could share the devices and handle multiple inputs simultaneously. In contrast,

Schwabe et al. (2005) cautioned that technology should not be the center of the learning

activity. Schwabe et al. contended that if the technology itself is at the center of the

learning experience, then the more likely it is for the technology to become a distraction

instead of an enhancement. Potential issues to guard against are student misuse of the

technology, technology failures, or the technology detracting attention from the learning

activity. Moreover, the current form of mobile devices, on board with AR technology,

allows users to be aware of their environment and others in it, which could increase the

opportunity for users to work cooperatively. Although collaborative learning is not new,

recent technological innovations have further expanded how cooperative work can be

accomplished in the classroom. The pervasiveness of powerful new mobile devices with

unlimited wireless connectivity (in different shapes and sizes) can potentially expand

how collaboration can be accomplished in the classroom. Therefore, learning

collaboratively in the context of emerging technological landscapes, such as mixed

realities, needs to be better understood (Dillenbourg et al., 2009).

Use of Emerging Technology to Support Statistics Education

Now more than ever, our lives are interconnected with technology. But what role

does technology have in the teaching and learning of statistics education, particularly for

teaching statistical reasoning? The ever-changing nature of technology makes it difficult to judge the specific impact of technology on statistics education (Kaput, 1992). However, it seems that pedagogic developments have struggled to keep pace with the affordances of new technologies (Pratt et al., 2011). The opportunity to use technology in the classroom to engage students to its full capacity has not yet been realized.

Nevertheless, technology has had a great impact on how statistics is learned, perhaps more so than many other disciplines (Chance et al., 2007). Through innovations such as computers, graphing calculators, software, and the Internet, the way statistics is taught has changed dramatically. In today's classrooms across all grade levels, you will most likely find a computer projected onto a screen with students working on powerful scientific calculators and perhaps even working at their own computers. Additionally, it is commonplace for students to complete statistics assignments on computers in a school lab or at home. In turn, some of the instruction is done via the Internet in the form of "Web-based courses with videotaped lectures, interactive discussions, collaborative projects, and electronic text and assessment materials" (Chance et al., 2007, p. 1).

Moving forward, it is apparent that technology has the potential to help students conceptualize and understand statistics beyond just serving as a shortcut to calculating values and outputting results. Computer-mediated instruction can help students learn basic statistics concepts by providing different ways to represent the same data set (e.g., going from tables of data to histograms) or by allowing students to manipulate different aspects of a particular representation in exploring a data set (e.g., changing the shape of a histogram to see what happens to the relative values of probability and variability). Statistics software packages may also be used to help students better understand abstract

ideas. For example, students may develop an understanding of normal distributions by constructing various samples and observing the distributions of statistics computed from the samples drawn from prescribed populations (Chance et al., 2007). The computer can also be used to improve students' understanding of probability by allowing them to construct their own statistical models, change assumptions and parameters for these models, and analyze the data generated by applying these models (Garfield & Ben-Zvi, 2007). Innovative visualization software, such as Fathom, is available to students at all levels to analyze data and possibly learn to reason statistically (Key Curriculum, 2013).

Research on new advances in technology such as games and simulations have shown promise as teaching tools. For instance, Klopfer and Squire (2008) investigated the merits of four popular AR mobile games that were designed for both education and entertainment purposes. The focus of their work was primarily on the user interface, interactivity design, and a framework for designing educational games and simulations as opposed to conducting an evaluation of these approaches. The researchers documented their development sequences during five classroom case studies. Klopfer and Squire (2008) concluded that as long as the interaction is well designed and carefully structured, games and simulations are appropriate for teaching topics such as social sciences, among others.

Researchers have explored the value of games and simulations for teaching basic statistical concepts. Lane and Tang (2000) compared the effectiveness of simulations to the effectiveness of a textbook for teaching statistical processes and found positive results. This study was conducted with 115 undergraduate students where the dependent variable was the students' ability to answer questions to everyday statistical problems. A

multimedia simulation developed by David Lane (1999) called the Rice Virtual Lab

showed the process of random sampling problems and contained a histogram of the

population distribution on the screen. Each time a random sample was drawn, the scores

were displayed next to the population histogram. Students were then shown the mean

scores. The process was repeated to show how sampling and sample sizes are conducted.

Students who studied using the simulation performed better than those who studied with

the textbook. However, the authors found that the simulation limited the students' ability

to interact with the tool. Specifically, the simulation was very limited, because the actual

interactivity (or manipulation of the data) happened outside of the simulation itself.

Engagement and interactivity are particularly important to today's students who

have grown up with high definition and extremely realistic computer games. Their

constant exposure to these types of games and other digital media found on the Internet

has inevitably shaped how they receive and process information to learn (Tan, Lewis,

Avis, & Withers, 2008). Although there is continuous debate on the use of games and

simulations in education, researchers continue to explore their potential to enhance

engagement and learning (Gee, 2007; Gee, 2003; Shaffer, Squire, Halverson, & Gee,

2005; Squire, 2005). There are many attributes of games and simulations that make them

pedagogically sound and engaging learning environments (Gee, 2003). A summary of

Gee's (2003) principles are: the environment is designed to encourage active, not passive,

learning; students need to be able to take risks in the environment where real-world

consequences are minimized; intrinsic rewards are customized to each student's level,

effort, and growing mastery; the environment needs to provide multiple practice

opportunities in a compelling context that is not boring, which may increase time on task;

and knowledge and meaning is constructed through various multimedia such as text, images, symbols, and animations. Correspondingly, a number of researchers have used games and simulations as enhancements to the traditional learning environment with encouraging results (Squire & Jenkins, 2003). However, more empirical evidence for and against games and simulations as a tool for learning is needed before completely signing off on its value. Gee (2011) went as far to state that the empirical results to date are a "mixed bag" (p. 224). While the engagement and interactivity of games and simulations are highly positive, a number of questions remain about how they are developed, deployed, and accepted by educational decision makers.

Beyond games and simulations, there are other ways emerging technologies offer the potential to expand the tool kit of instructional methods that educators can use to support statistics education. Due to the ubiquity of powerful mobile devices and the extension of the classroom into the real world, it seems sensible to consider pedagogical and technological ways to blend virtual concepts with the physical environments as a teaching tool (Dede, 2005).

Augmented reality. Azuma (1997) defined AR as "an environment that includes both virtual reality and real-world elements" (p. 357). Azuma et al. (2001) later expounded on the initial definition of AR to include the properties of combining real and virtual objects in a real environment and running interactively in real time. Different researchers subscribe to variations of this definition (Dede, 2009; Höllerer & Feiner, 2004; Klopfer & Squire, 2008). Nonetheless, the research community largely agrees on the defining elements of AR systems, but there are just small differences in the

16

incorporation of new technology. For the purpose of this study, the definition from Azuma et al. (2001) was used.

The technology, hardware, and software behind AR have matured to the point where it can be more readily deployed for educational purposes in varying physical environments (Kroeker, 2010). Despite its considerable underlying sophistication, most educationally oriented AR strives to be relatively approachable for the average classroom teacher. Usability is a key consideration for teachers adopting technology into his or her instruction (Christian, 2006). It is important that the technology be easy enough to use so teachers are not intimidated or apprehensive about utilizing new technology in their classrooms, or they may not adopt it. In reality, we already have the ability to integrate AR technology using typical mobile devices, such as smartphones and tablets. To generate an AR experience, only three key components are required: a camera with a display screen, a computer processor, and the appropriate software. Today, almost everyone already has a device with the combination of these technologies in his or her possession (Christian, 2006; International Telecommunications Union, 2012).

Wagner and Schmalstieg (2003) described one of the first stand-alone AR systems that was created using just a commercial camera. Using readily available consumer devices, the AR application was created with popular marker-based tracking software while running on a basic wireless network. This work showcased the minimal need for highly specialized or costly technologies to create AR experiences. Katz, Cook, and Smart (2011) also noted that the programming skills needed to create AR scenarios are as straightforward as those found in most basic Web design tools. Common technology and AR development software is now at a point where the process of

generating AR scenarios is as intuitive as possible without needing much coding or advanced technical expertise. AR applications are increasingly easily created and they are also powerful. Not only is it easy to create AR scenarios, but most current versions of mobile devices also come equipped with the necessary software for creating and running AR scenarios. This accommodation is conducive for leveraging AR for educational applications. Now more than ever, students are exposed to and possess these emerging technologies, therefore becoming more intuitive for them to use (Höllerer & Feiner, 2004).

AR has existed in the mainstream for some time, but has not yet been fully embraced in the education sector (Kroeker, 2010). Historically, it was considered too complex and expensive, and the supply of educationally relevant content was very limited (Höllerer & Feiner, 2004). It consequently has drawn little attention as an educational technology until the last decade. Most notably, Klopfer, Squire, and Jenkins (2002) established a research program around an AR game designed for learning called Environmental Detectives. This game was a multiplayer, handheld AR simulation intended to be used in high school and undergraduate settings. The purpose of the game was for teaching learners environmental inquiry and reasoning skills. Given a contextualized role of being environmental engineers, the learners (while working in pairs) had to diagnose the root cause of an environmental problem and resolve a toxic spill threat to the local ground water source. The functionality of the game was to determine contamination levels by sampling the drinking water and to collaborate with virtual experts. The activities of the game were conducted in a defined physical space using a global position system (GPS) enabled handheld computer. Klopfer et al. (2002)

concluded that the individual learner's experience was active even though it was mostly simulated. Because the learners collected the location-based data in real time, they were unable to discern if the information was real or simulated. This is an affordance provided by the natural way a mobile device and AR complement one another for educational purposes.

More recently, another team of researchers explored the affordances of AR for data collection in a similar context. The Ecosystems Mobile Outdoor Blended Immersive Learning Environment (EcoMOBILE) project was designed to augment and supplement an elementary school environmental science field trip (Kamarainen et al., 2013). While visiting a local pond, a group of sixth graders collected data in order to identify different biological qualities found in the ecosystem. The students used smartphones to view virtual information from strategically placed AR targets positioned around the field trip location. Using the smartphones, the students were able to view information that normally would not be available without some form of technology. At the conclusion of the field trip, the researchers measured engagement and collaboration. According to these measures, they documented positive benefits of using AR and mobile technology. The authors felt that their work provided an example of how technology can be harnessed for educational purposes to create a learning experience that is student-centered and also provided opportunities for collaboration. Additionally, the researchers recommended that an AR curriculum should include both real and simulated experiences; the technology should be as authentic as possible with instructional cues and navigation guidance. They concluded that their ability to design and recreate an authentic, immersive environment was promising, yet required further exploration needed.

Engagement and learning gains are often reported in mixed reality environment studies such as Taiga Park (Barab, Sadler, Heiselt, Hickey & Zuiker, 2010), Quest Atlantis (Barab, Thomas, Dodge, Carteaux, & Tuzun, 2005), Whyville (Kafai, Quintero, & Feldon, 2010) and Alien Contact (Dede, 2009). For instance, Dede's (2009) AR investigative simulation Alien Contact was used to develop complex problem-solving skills in math, language arts, and scientific literacy skills for middle and high school students. The AR simulation was designed to increase engagement by including a narrative, a setting, role-playing, a master goal divided into subtasks, interactivity, choice, and collaboration (Dunleavy, Dede, & Mitchell, 2009). To measure engagement, the researchers monitored through observation examples of student engagement such as when students appeared to lose track of the real environment and focused completely on the AR simulation. A major finding reported in this study was that AR and mobile technologies provide unique opportunities to create authentic and novel learning environments. The findings indicated that the actionable and symbolic metaphors increased engagement regardless of demographic variables such as gender, ethnicity, or English language proficiency. The researchers concluded that students were more engaged in the immersive experience and showed improved learning gains overall. Similarly, creators and researchers of the Taiga Park study, Barab et al. (2010), posited that in these emerging learning environments, "the educational promise of virtual reality lies not in engagement with the media, but with the narrative; not only through sensory immersion, but also through narrative immersion" (p. 403).

In general, it is worth noting that there previously has been little consensus on the definition of *engagement* in the literature (O'Brien & Toms, 2010). Likewise, there is

little agreement on how it should be scientifically measured (Chapman, Selvarajah, & Webster, 1999; Coller, Shernoff, & Strati, 2011; Nijholt & Vinciarelli, 2012; Witmer & Singer, 1998). However, over the past 15 years there has been a surge of effort devoted to defining engagement across multimedia disciplines such as advertising, retail, graphic design, entertainment, and other visually rich contexts (O'Brien & Toms, 2010). Consequently, a variety of data collection techniques have been established to measure engagement. In educational technology, the most commonly used techniques are self-report measures, including Keller's (1987) Instructional Materials Motivation Survey (IMMS) 20-item survey that focuses on attention, relevance, confidence, and satisfaction and Webster and Ho's (1997) seven-item questionnaire with items pertaining to attention, challenge, intrinsic interest, and variety. However, these self-report survey instruments designed to measure engagement have yet to be applied in domains such as AR or mobile technology. While self-report measures are not as objective as recording actual performance or measuring engagement with physiological indicators, they do offer a convenient and efficient means of assessing the users' perspective of an experience (O'Brien & Toms, 2010).

As a commonly accepted approach, a self-report survey instrument was adapted and used for this study to collect participants' perceptions of their level of engagement during the learning activity. Specifically, due to its ability to take into consideration the multifaceted nature of engagement, a tool created and validated for online shopping experiences by O'Brien and Toms (2010) was adapted and used. As suggested by O'Brien and Toms, this self-report instrument is an acceptable and robust tool for assessing users' responses to an existing technological system. The instrument not only

provides a general evaluation of the users' experiences, but it also captures users' perceptions of the attributes such as focused attention, aesthetics, perceived usability, endurability, novelty, and involvement. As a result, the O'Brien and Toms engagement scale was chosen for this study because it includes more affective attributes that can be related to the technological aspects of this study more greatly than instruments used in other studies. For instance, if students overwhelmingly rated the aesthetics during the AR experience as low, it would be easier to pinpoint any confounding issues with that attribute or others.

Mobile technology. Wireless, portable computing technologies such as laptops, tablets, and smartphones have become an integral part of our daily lives (So, Kim, & Looi, 2008). Experts have coined this time in our history as the digital age of the 21st century, a phrase that reflects the paradigm shift in how we access and consume information provided by and accessible on computing devices (Looi et al., 2010; Quinn, 2011). With currently over 6 billion mobile device users globally (International Telecommunications Union, 2012), this revolution will undoubtedly have implications for future user behaviors, including what it means to be a student. Consequently, the advances in mobile technology have prompted educational technologists to consider new ways to transform and extend the learning environments to leverage mobile devices to facilitate anytime, anywhere, untethered learning (Naismith, Lonsdale, Vavoula, & Sharples, 2004; Quinn, 2011; So et al., 2008).

The Mobile Augmented Reality System (MARS) project was one of the first of it's kind to attempt to investigate the value of AR technology (Höllerer, Feiner, Terauchi, Rashid, & Hallaway, 1999). The researchers behind the project created a mobile AR

22

system that turned the outside world into a classroom by allowing the students to freely walk around outdoors while having all necessary equipment mounted onto his or her back to view virtual images on real-world objects. MARS consisted of two different experiences that permitted students positioned indoors to view digital objects transmitted to them from students positioned outdoors. The outdoor group of students walked around outside equipped with a head-mounted heads up display (HUD) connected to a hand-held mobile computer. The inside group viewed visual data that merged virtual with real-world information through a desktop while sitting indoors. Through direct communication via the mobile device, the students indoors guided and directed the students outdoors to change what was displayed. Based on this concept, Sharples (2000) also designed and evaluated a portable computing system used to support learning from any location. Mediated by life-long learning theory, the system was designed to allow elementary students the ability to communicate with each other, as well as with the teacher. The system was one of the first to show how to conduct situated learning activities with mobile devices. Höllerer et al. (1999) and Sharpless (2000) are the two studies that provided the initial framework for the software, hardware, communications, and interface design for mobile devices as a learning resource that served as the foundation for the novel technology at the center of this study.

Schwabe and Göth (2005) extended the framework with the design of a mobile learning game for undergraduate students that provided a mixed reality that augmented an indoor and outdoor physical space. The authors found that students responded positively to features such as map-navigation and hunting and hiding. However, the functionality of the prototype brought up design issues in the accuracy of GPS due to connectivity

limitations. The conclusion was that the system required a stronger Internet connection for better real-time response rates for a smooth and seamless experience as the students moved throughout the physical space.

Kurti, Milrad, and Spikol (2007) also developed a handheld mobile system for both an outside and classroom learning experience for elementary students. This was one of the first studies to suggest the notion of ubiquitous computing to support contextualized learning activities. The purpose of this study was to enhance the content of the curricula by creating a contextualized, collaborative learning game where fifth grade students had to complete missions. The context of the game was that students completed a quest as a blacksmith from past centuries solving navigational clues. Although many of the missions were historically based, students in some cases had to solve math problems to understand clues. The game was played throughout a small town in Scandinavia using a smartphone, where onboard mobile device features such as wireless Internet connectivity, computer processing, and instant messaging were used to complete the tasks. The researchers, however, suggested further effort was required on the interface design on the mobile device to better fit the learning activity. In turn, this would increase the authenticity of the experience for the students. The overall conclusion was that the technology successfully functioned for the purposes of the study and that the students responded positively to the experience.

Costabile et al. (2008) conducted an experimental pilot study of a mobile system called Explore! as a learning companion for middle school students on a history field trip. The students, working in small groups of three to six people, answered questions and navigated an archaeological park using a cell phone to solve a mission. The goal of the

study was to make the visit more engaging, and to make the information more meaningful and relevant for the students. The researchers also assessed the reliability of the mobile system by comparing the students' experience with and without technological support. The researchers examined data such as time on task, behavioral observations, and questionnaire information. The researchers found that students with and without technical support both had positive responses to the activity. They also concluded that students did not have difficulty using the mobile devices even without technical support, and that it empirically had pedagogical value for providing guidance during a learning activity.

Fotouhi-Ghazvini, Earnshaw, Robison, and Excell (2009) conducted another study designed to examine the educational application of mobile technologies by comparing and evaluating four AR games. To answer their research questions, several physiological sensors from a popular smartphone were used to collect data from the players' point of view (i.e., GPS, radio-frequency identification, Bluetooth, infrared, and camera). The information gathered was aligned to establish a design protocol and implementation strategy for using AR games in the classroom setting. Using a widely accepted instrument to measure engagement, Keller's (1987) IMMS, the players' perception of attention, relevance, confidence, and satisfaction for the instruction from each game was measured. The researchers concluded that AR added a real sense of learning to mobile games by creating concrete connections of abstract concepts for the students. The researchers proposed the following set of best practices for effectively integrating the emerging technologies of AR with mobile devices:

- Create virtual experiences that mirror the real world

- Constrain extraneous information

- Provide appropriate feedback, and

- Provide context throughout the experience

The form factor and mobility of this AR system facilitated in-class cooperation in small groups better than other technologies. Technologies such as desktop or laptop computers do not allow for easy sharing of input devices (Billinghurst & Kato, 2002). For instance, observing a computer screen or swapping turns sitting at the controls can lead to a passive mode of participation for students (Höllerer & Feiner, 2004; Schwabe et al., 2005). On the other hand, AR on a tablet creates a more natural face-to-face collaboration, where students can use speech, gesture, gaze, and even nonverbal cues to communicate to solve problems (Billinghurst & Kato, 2002). Lastly, this AR system potentially bridges the separation between the real world and the virtual 3-D objects used for the learning task (Billinghurst & Kato, 2002). This affordance allows students to collaborate on the appearance of objects such as size and color and use metaphors or other representations as well as facilitates the ability to establish a point of reference for communication. This last benefit is particularly important due to the nature of the planned activity to be used in this study. In this study, some students were assigned to work in pairs to complete the learning activity while moving around a classroom.

A natural extension of mobile-enabled AR experiences is providing opportunities to embed learning in authentic, collaborative environments, which potentially enhances engagement and learning outside traditional formal educational settings (Huizenga et al., 2009). As such, the ubiquity of mobile technologies has created an opening where new possibilities for mobile devices and AR are conceivable. Students are no longer tethered

to a classroom or library computer to discover, explore, and learn. By way of mobile devices, students are able to obtain information, collaborate, communicate, and complete educational activities for the purpose of pursuing knowledge while governing their own learning. Previous mobile learning research, however, has typically focused on either mobile technology or AR separately, without examining the integrated and synergetic effects of linking these two emergent technologies. Furthermore, there has been very little experimental research conducted to investigate the potential impacts of these emerging technologies on statistics education.

Embodied Cognition

There are many different theories that explain how the body is connected to the mind. Embodied cognition is the primary learning theory that encompasses most of these theories as the body and mind work together for learning purposes (Wilson, 2002). It provides one of the original theoretical explanations of how the body subjugates particular bodily systems such as perception, action, and emotion for higher cognitive processes (Varela, Thompson, & Rosch, 1991). Furthermore, embodied cognition helps explain how these processes are important to learning, such as language comprehension, reading, mathematics, and scientific thinking (Glenberg, 2008). As such, statistics education is ideal for observing the effects of embodiment and investigating these theories, and as such, will be included in this study.

The unification of the mind and body for the purposes of learning begins at birth (Glenberg, 2008). As infants develop their motor skills, they create different schemas by exploring the world around them through the use of their senses (Piaget, 1953). Within the classroom, students similarly develop a spatial context of information anchored to

their recollection of certain memories, emotions, or actions that have taken place. Embodied cognition, in the context of this study, describes how students will use their bodies and movements to complete a learning task, which is possibly closely tied to the development of a deeper understanding of statistical reasoning concepts. Many educators already use embodied cognition to ingrain their instruction in this area. Through the use of images, classroom discussions, and narratives, teachers often try to connect objects or metaphors with concepts to help students learn. Although there is a considerable amount of research on the topic of embodiment in education, clear educational applications for statistics education has yet to be agreed upon.

Price and Rogers (2004) offered a framework that defines three key elements for adequately creating digitally augmented physical spaces: (a) the interaction with physical tools, (b) physical movements, and (c) combining objects with each other. The interactions with physical tools component includes actions like taking notes with a pen and paper, writing on a whiteboard, or solving a problem using a computer. These actions are customary ways that help students demonstrate their cognition.

A consideration of this framework in terms of this study involves how extending students' cognitive actions with physical tools can be further reinforced by being coupled with new forms of physical tools such as mobile devices. With the help of tablets and smartphones, students can make-believe that they are using a specialized device for the purposes of conducting an investigation of an ecosystem (i.e., microscopes, thermometers, or calculators). Price and Rogers (2004) contended that through "manipulation, scanning, and motion," students are more likely to activate different cognitive resources (p. 141). Thus, providing students with the opportunity to see

different perspectives of physical objects can provide the opportunity for deeper cognitive connections. Having students get out of their seats to move around the classroom could be enough to stimulate other senses more than just passively listening to instruction. Remembering the feel of a mobile device in his or her hands could also trigger recall of important steps of an activity during a test. Furthermore, students could also conceivably see and interact with 3-D objects projected on mobile devices as tools such as "hammers, spades, or wands" to simulate certain activities (Price & Rogers, 2004, p. 141). Price and Rogers' (2004) key element of physical movement describes how mental models are formed in relationship with the positioning and moving the body in space.

Combining objects with each other is probably the most significant principle of AR. Integrating two objects beside, inside, or on top of one another is a unique affordance of AR, which promotes placing practice closer to the dissemination of instructional information—a major premise of this study (Price & Rogers, 2004). For example, showing students the combination of artifacts such as virtual images of a fish swimming over an image body of water could create a sense of exploration. Given the goal of determining how many fish have a particular characteristic in a group offers the opportunity to practice a skill like sampling. This method is an alternative either to using real fish in a real pond or paper fish on a real pond. While both situations provide challenges, finding and working with real fish is impractical from a learning perspective. Using paper images of fish are less interactive and could by some degree reduce the authenticity of the experience. Price and Rogers' (2004) research in this area has shown

that working with digital information in physical spaces can increase interactivity and engagement.

A modest number of studies have examined embodiment or a related framework for making the association between embodiment and learning performance in the classroom. The most notable, Cook, Mitchell, and Goldin-Meadow (2008), investigated the impact of physical movement in the form of gestures. In this experiment, fourth grade students learned to solve new mathematical concepts via one of the following conditions: speech, a gesture, or speech with a gesture. Students assigned to the speech condition had to solve a problem by verbalizing to the teacher how to solve a basic equation problem. Students assigned to the gesture condition had to solve a problem by demonstrating a problem-relevant gesture (i.e., sweeping the hand under the left side and then the right side to show that one side is equal to the other in an equation). Students assigned to the speech with gesture condition had to solve a problem using both techniques at the same time. Four weeks later, the students were tested again, and those who had been taught the gesture showed significantly more retention from the initial instruction. The authors concluded that an embodied representation of new mathematical concepts through gestures could have an impact on learning gains.

In a study by Chan and Black (2006), researchers created a specialized digital widget module called direct-manipulation animation (DMA) to help middle school students visualize basic physics concepts. Each student was randomly assigned to one of the DMA conditions (narrative only, narrative and static visuals, and narrative and animation) for a physics lesson. Depending on the assigned condition, students were able to examine a computer-based widget with varying levels of navigation control to see

causal interactions and functional relationships. Using a pretest and posttest design, the researchers measured students' understanding, retention, and near and far transfer of mechanical energy concepts and ideas. Additionally, students' perceptions from each condition were collected using an online survey. Chan and Black found that interacting freely with the navigation controls to manipulate the widget in the narrative-and-animation condition provided learners with a superior learning experience, as compared to those who received narrative-only and static visuals about gravity and forces. The authors described that as the content became increasingly complicated, DMA proved to be an effective support to assist the students in comprehending, reasoning, and solving learning tasks. This is relevant for the first factor of the embodiment framework used in this study. The interaction with physical tools potentially aided the learning process from the immediate cause and effect response from hand gestures. The researchers suggested a modification to this study would be to examine how smaller devices could be used to deliver the content and how they could impact embodied learning experiences.

Enyedy, Danish, Delacruz, and Kumar (2012) also attempted to gain further insight into an embodied understanding of space in a study of young elementary students on a related scientific investigation of Newtonian force and motion. Using a collaborative AR system, students' physical actions were recorded via a Web camera on a computer located within the physical classroom. The AR system displayed the concept of force or motion and the students responded by duplicating the concept, which was converted by specialized software based on the sensing data. They compared pretest and posttest results and showed that the elementary students were able to develop a deeper conceptual understanding of force, net force, friction, and 2-D motion after a sequence of AR

31

activities. The authors also found that physical movement resulted in more authentic symbolic representations of the curriculum in the learning activity. Additionally, in one of the case studies presented by Enyedy et al., they suggested that embodied experiences provide a safe place to share ideas without negative consequences. Enyedy et al. offered a proposed design framework for implementing embodied experiences in the classroom similar to the one selected for this study.

This body of research continues to show how digitally augmented physical spaces can be designed to exploit interactional capabilities, which are enabled by wireless networks and pervasive technologies, to support learning in quite different ways than have traditionally been the case with other fixed location, computer-mediated interactions. Still, it is thought that the field of statistics education has not kept pace with the credence that the related literature gives to the importance of embodiment on young students' cognitive capabilities that can be used in concert for understanding statistical concepts (Núñez, 2012). The inclusion of mobile AR technologies in traditional classroom spaces provides untapped opportunities for a new genre of physical–digital interactions that can support active learning, and, in particular, situation, exploration, and cooperation (Price & Rogers, 2004).

AR researchers such as Price and Rogers (2004) contended, "one of the key aspects of interacting in digitally enhanced physical spaces is to raise the awareness of the children as to what they are doing in them" (p. 148). Another core aspect is that AR can provide a richer experience (compared with virtual worlds), allowing students to make explicit connections between their various perspectives and understandings of the physical and digital worlds both at the same time. Other benefits are that engagement

could be triggered when pairing the familiar physical actions and unfamiliar effects in order to solve scientific problems (Anderson, 2003). The degree of authenticity of the learning experience and the amount of collaboration that results—both of which are considered in the literature to be important aspects of embodied cognition—can also be greater with using AR (Price & Rogers, 2004).

What further makes the consideration of embodiment appropriate for this study is the exploration of AR as interactive in real time that can be measured and correlated to specific student positions, movements, and actions, thereby making it entirely consistent with the focus on embodied experience. Based on the literature, there is a need to expand the impact of interaction and real time in relationship to the technology and students. Furthermore, because AR is being evaluated for how it allows students to move around a physical space, this scientific approach is true to the mobile nature of the application of these technologies in a classroom setting. As such, the focus of this study is to explore and define a system that enables students across all levels to interact with the physical world while viewing relevant augmented digital information, which can subsequently be interacted with and facilitate a more active way of learning.

Situated Learning

Anchored in the embodied cognition area of literature is the assertion that "cognition is a situated activity" (Wilson, 2002, p. 626). As originally defined by Lave and Wenger (1991), situated learning is the pedagogical theory that cognitive activity is situated in a specific context and embedded within a particular social and physical environment. Additionally, Wilson (2002) stated that "cognitive activity takes place in the context of a real-world environment, and it inherently involves perception and action"

33

(p. 626). For instance, Glenberg (2010) reasoned that cognition is the synthesis of abstract symbols such as words, numbers, and symbols grounded in an embodied experience. Glenberg tested this theory by developing an intervention, called Moved by Reading, to help young children develop their reading comprehension skills. There were two treatment conditions: physical manipulation and control. In the physical manipulation condition, children read a contextualized scenario, such as what a farmer does on a farm, while interacting with representative toys (e.g., a toy barn, corral, tractor, animals) that corresponded to parts of the text. Children in the control condition read the same passages, but were not given the opportunity to manipulate the toys. Glenberg's (2010) findings were that the physical manipulation of the toys improved comprehension. He concluded that connecting symbols and their embodied meaning was an effective way to help children make connections on their own.

As characterized previously in this chapter, AR naturally affords situated learning. AR provides learners the opportunity to connect symbols and their embodied meanings by entrenching cognition in both a symbolic and physical context, distributed between learners and the tools used (Kamarainen et al., 2013). Barab and Dede (2007) suggested that learning science "should be situated as an inquiry process and that new technologies and design methodologies can facilitate this process" (p. 1). However, for exploratory and situated learning in physical contexts, the focus should not be limited to the device (Schwabe et al., 2005). All tools and resources should be situated to produce a learning synergy that scaffolds and supports scientific thinking (Squire, 2008). In the framework of this study, mobile devices with an AR application were selected as the technological tools to enhance the situated learning experience.

Dede (2009) also stated that the context is important not only for creating collaborative, embodied learning experiences, but to also help facilitate transfer (which is a major void in statistics education). A developing subset of this area of the literature is a new theory called immersion. This concept of immersion is "the subjective impression that one is participating in a comprehensive, realistic experience" (Dede, 2009, p. 66). Immersion in a digital experience involves the willing suspension of disbelief and the design of immersive learning experiences, which draws on sensory and symbolic factors. Sensory immersion is created by the replication of a digital location inside a 3-D space. Through interfaces such as head-mounted displays or immersive virtual rooms, total sensory immersion is a deeper form of immersion. Suspension of disbelief of reality is created by stereoscopic sound and, through haptic technologies that apply forces, vibrations, and motions to the user, the ability to touch virtual objects. As described, interactive media now enables various degrees of sensory immersion (Dede, 2009).

The thought behind situated investigation and sensory immersion in a real-world context is that it may facilitate transfer. Specifically, applying a contextualized learning experience may enable preparation for future learning in that students learn skills that may be applicable to learning more generally (Anderson, Reder, & Simon, 1996). Considerable effort can be expended in trying to help students transfer their knowledge from the classroom to the real world. Bringing AR technology into the classroom could supplement or replace the need for cumbersome or expensive field trips. Although there are many benefits of real-world field trips, an augmented, situated learning experience can also provide safe places for students to interact with virtual objects without adverse consequences. Additionally, the instructional focus on transfer can be more readily

35

controlled and applied to other real world contexts (Kamarainen et al., 2013). As such, another focus of this study was on exploring the unique affordances of AR that can support this kind of situated learning in statistics education. Notwithstanding, the learning environment is interdependent to the learning process, because the context can alter, enhance, and support certain types of performances, approaches to problems, or learning activities (Kamarainen et al., 2013; Squire & Jan, 2007).

Overview of Present Study

Purpose. This study was designed to explore whether AR learning experiences can support learning. The study is motivated by an interest in exploring educational applications of AR that could potentially provide engaging contextualized learning for understanding concepts related to statistics education. The hope is that this research contributes to the development of empirical strategies for leveraging AR with learning complex skill sets. This study was designed to help researchers and practitioners better understand how learners build knowledge and engage in context-based, collaborative learning experience using AR combined with mobile technology.

Importance. Initial evidence points to the potential use of AR technology as a distinctive tool for creating engaging, collaborative learning opportunities with spatially situated learning experiences. Research suggests how this emerging technology, in conjunction with mobile technology, can be used to help students move about a physical space where contextually relevant information and resources are provided digitally, and in such a way that students can collaboratively explore, capture, and manipulate both physical and virtual objects for active understanding.

AR offers a distinct set of affordances for statistics education. As noted, AR coupled with mobile technology, unlike other technologies, permits students to interact with 3-D objects while moving around a physical space (Starner et al., 1997). It also provides them with an overlay, onto everyday objects, of helpful information not normally seen by the naked eye. Due to the form factor of mobile devices such as smartphones and tablets, the technology can be easily shared in such a way that provides opportunities for collaboration worth considering for classroom activities.

Given the advances in mobile (and wireless) technologies, it is now feasible to extend the standard classroom to create embodied learning experiences. With the addition of AR as a natural complement to mobile computing, students can interact with virtual objects or data in the real world in ways previously not considered. The compelling part of these technologies is that together they can deliver a unique and engaging learning experience that inherently creates unique collaboration opportunities. Additionally, the mobility will allow the students to move around, helping them anchor their understanding in the mind and body. As Enyedy et al. (2012) suggested, "it may be that the embodiment gives a physical sense of the extreme nature of this change that is not conveyed in symbolic representations" (p. 369). However, because of the novelty of AR technology, there is no off-the-shelf methodology for statistic education experts or educational technologists to follow.

The key issues raised in this study are how we can effectively leverage the new technologies pedagogically and what the impacts are. Even though AR is beginning to permeate our lives through smartphone apps and viral videos from technology firms like Apple and Google, existing uses are still mostly novel. Admittedly, we have further to go

37

before mobile AR technology has a real effect on everyday learning. In order to exploit the positive characteristics of AR technology for learning, the research in this area suggests that existing frameworks need to be extended and new frameworks must be defined, hence the motivation behind this experimental study.

Learning activity. The over-arching learning activity in this study was to calculate the proportion of fish in a body of water that had been affected by an oil spill. According to the assigned AR treatment group (high AR, low AR, or control), which will be defined later, participants viewed and interacted with strategically placed 2-D quick reference (QR) codes located in a classroom, which were readable through the live-camera view on a mobile device. Given the contextualized learning activity, the QR codes were used to project the renderings of fish in a pond. To complete the activity, the participants identified and approached the QR codes in sequence using the mobile device wirelessly connected to the Internet. Using the onboard technology, students received interactive guiding prompts to answer questions related to the contextualized problem.

One of the first questions for the participants was to estimate the percentage of fish infected by the oil spill out of a random sample of fish. To identify which fish were affected, students had to look for an oil smudge on the scales of the fish. The number of fish infected in each sample varied between three different samples taken by the participants. Using simple math, participants then used the calculated sampling of fish to estimate a sample proportion of how many fish were infected. Next, the participants were asked to compare their results with different samples taken by other participants. Participants also were prompted to analyze sample proportions of fish infected by pollution in the different areas of the body of water. During the last stage of the scenario,

participants compared their results to the actual solutions to make statistical inferences about variance between samples.

Consequently, Apple iPad mobile devices preloaded with an AR application designed by the researcher were used for this study. The underlying learning activity was designed to help learners better understand proportions, sampling distributions, and sampling variability, given a particular premise. The context for this study was that the participants would calculate the proportion of fish that were infected by pollution in a collection of ponds. By analyzing fish of varying species via 3-D virtual images or 2-D paper-based images, participants estimated sample size, statistical proportions, and identified factors affecting sampling variability. By following the instructions and prompts embedded within the AR experience, participants were guided through the collection of fish affected by pollution while given the opportunity to reflect on the calculations to make sense of the data. This was followed by feedback that clarified or reinforced the concepts.

Conditions. Using a 3 x 2 factorial research design, this study compared and analyzed the performance of students given two factors: AR experience and collaboration. The levels of AR experience were a high AR (3-D images with movement throughout a physical space) experience, a low AR (3-D images without movement throughout a physical space) experience, and a no AR (traditional 2-D images and text) experience to serve as the control; the levels of collaboration were either completing the learning activity in pairs or individually. As suggested by the literature, the AR application was designed to provide hands-on experience working with statistical data in

a way that students could control, modify, and repeat, all while in a natural classroom setting (Moore, 1997).

Despite a lack of consensus regarding ideal group size, participant performance in pairs was investigated in this study. The data collected from participants who worked in pairs versus those who worked individually helped provide answers involving the benefits of collaborative learning. Working in pairs was chosen because assigning more than two participants to a group could have possibly confounded the results of the study if someone was less active than desired. As group size becomes larger, the potential for an individual to contribute decreases (Roseth et al., 2008). Also, the smaller the groups, the easier it is for a supervisor such as a teacher or researcher to recognize and intervene to remedy any difficulties participants may experience during the collaborative learning activity. This is especially important with short activities and when participants are not familiar with the task (Roseth et al., 2008).

Research questions. This study investigated the effects of varying levels of a mobile AR system for learning complex statistical concepts. Furthermore, the impact of varying levels of collaboration was analyzed. Specifically, the research questions were:

1. Does an AR experience combined with collaboration impact learning gains in participants' understanding of statistical reasoning concepts beyond a no AR experience? Specifically:

   a. Does an AR experience (high AR, low AR, or no AR) produce differences in learning gains?

   b. Does working individually versus collaboratively produce differences in learning gains?

40

c. Do different levels of AR experience and prior knowledge interact to produce differences in learning gains?

2. Does a high or low AR experience facilitate the perception of a more collaborative experience than a learning experience without AR?

3. Does a high or low AR experience facilitate the perception of a more engaging experience than a learning experience without AR?

Hypotheses. Based on the theoretical and empirical justification, the following hypotheses were investigated:

- H_1: Participants in the high and low AR experience will have higher learning gains than participants in the no AR experience.

  H_1 (Learning Gains): AR (High or Low) > Control

- H_2: Participants in the high AR experience will have higher learning gains than participants in the low AR experience.

  H_2 (Learning Gains): High AR > Low AR

- H_3: Participants in pairs will have higher learning gains than participants who work individually.

  H_3 (Learning Gains): Pairs > No Pairs

- H_4: Participants who work in pairs in the high and low AR experience will perceive more collaboration than participants in the no AR experience.

  H_4 (Perception of Collaboration): AR (High and Low) > Control

- H_5: Participants in the AR experience in the high or low AR experience will perceive more engagement than participants in the no AR experience.

  H_5 (Perception of Engagement): AR (High and Low) > Control

Variables

Independent variables

1. AR Experience

    a. High AR experience allowed the participants to complete the tasks with movement around the space.

    b. Low AR experience restricted the participants to complete the tasks without moving around the space.

    c. No AR experience was the control group and participants completed the same tasks except without the use of AR.

2. Collaboration

    a. Participants completed the learning activity in pairs.

    b. Participants completed the learning activity individually.

Moderating variable

1. Prior knowledge

    a. High prior knowledge was defined as the participant's familiarity with the content prior to participating in the study calculated by participant's pretest score one standard deviation above the overall pretest mean scores.

    b. Average prior knowledge was defined as the participant's familiarity with the content prior to participating in the study calculated by participant's pretest score equal to the overall pretest mean scores.

    c. Low prior knowledge was defined as the participant's familiarity with the content prior to participating in the study calculated by

participant's pretest score one standard deviation below the overall

pretest mean scores.

Dependent variables

1. Learning gains measured by participant's difference scores from pretest to
   posttest

2. Perceptions of collaboration measured by the participants' collaboration scale
   responses

3. Perceptions of engagement measured by the participants' engagement scale
   responses

Chapter 2

Method

Participants

Participants targeted for this study consisted of 252 undergraduate and graduate

students from two universities in the southwest region of the United States. From varying

programs at a major university, 115 undergraduate students and 79 graduate students

were selected to participate in the study. Also, a group of 58 graduate students from a

second nearby university were selected to participate in the study. In total, the students

who participated in this study represented over 35 different academic disciplines across

the varying class levels. A slight majority of female students (51% women, 49% men)

participated in the study. Participants ranged in age from 18 to 58 years with a median

age of 26 years. Other descriptive demographic information collected included ethnicity

and current education level and is summarized in Table 1. Because all of the participants

were over the age of 18 and no identifying information was collected, an Institutional

Review Board exemption status was granted for this study. Data from all participants

with a complete, matched pretest and posttest were included in the analysis. Only one

participant voluntarily withdrew, and two incomplete responses were removed from

inclusion in the final results.

Table 1

Sample Demographics by Condition

	Pairs			No Pairs		
	High AR (N = 39)	Low AR (N =53)	No AR (N = 41)	High AR (N = 41)	Low AR (N = 41)	No AR (N = 37)
Gender						
Men	49%	40%	58%	54%	49%	41%
Women	51%	60%	42%	46%	51%	59%
Median Age	26	28	28	27	25	24
Education						
Undergraduate	56%	53%	42%	54%	7%	62%
Graduate	44%	47%	58%	46%	93%	38%
Average of Statistics Courses Taken						
Zero	20%	26%	15%	29%	46%	38%
1-2	67%	68%	73%	56%	51%	51%
3-4	13%	6%	10%	15%	3%	11%
5+	0%	0%	2%	0%	0%	0%
Statistics Ability	4.05*	4.02*	4.15*	4.17*	4.02*	4.14*
Attitude Towards Statistics	5.23*	4.96*	4.71*	4.78*	4.68*	5.03*
Prior iPad Usage						
Never	8%	13%	7%	5%	14%	5%
A little	44%	55%	59%	46%	39%	51%
A lot	31%	19%	27%	37%	44%	43%
Other tablets	2%	9%	7%	12%	2%	0%
iPad Ability	6.87*	6.15*	6.59*	7.29*	7.12*	7.57*

Note. Participant self-reported mean scores reported on a scale of (1 = Poor, 10 = Excellent)

Learning Environment

The materials and equipment required for the design and execution of this study included a classroom, programming software, quick reference (QR) codes, mobile devices, and a learning activity.

Classroom. The study took place in a typical university classroom that had the capacity to host 20 to 40 participants and allowed enough space for multiple pairs of participants to move around without interference. The classroom included student desks and chairs as well as wireless connectivity.

Programming software. Qualcomm's Vuforia programming software was used to design and program the functionality of augmented reality (AR) components for the AR version of the learning activity. Vuforia is a software program with an easy-to-use application designed to create digital AR content. Using this programming software, AR content such as three-dimensional (3-D) images with text was created to deliver the AR version of the learning activity within the classroom. Additionally, the Unity 4 software package was used to design and develop the AR environment and interactions. Unity 4 is a game engine that seamlessly allows the programming of images, animations, and other special effects.

QR codes. To recognize and project virtual 3-D elements on real-world objects around the classroom, the AR scenario used personalized QR codes. In comparison to other ways of identifying an image or object, QR coding is an inexpensive method; therefore, QR codes are ideal for the purposes of this study. Customized QR codes were easily created and were a standard feature of the prescribed software application Vuforia. Vuforia software was also used to program the recognition of the markers for the AR

version of the learning activity. The AR experience leveraged three QR codes for the tracking and rendering of the various images of fish, gaining the camera perspective for correct augmentation of color and other visual hints. In the context of the fish theme, the QR codes were depicted as a body of water, as shown in Figures 1, 2, and 3. The QR codes were designed to resemble one of the three areas of a bay to support the water life theme of the narrative of the learning task.

Figure 1. QR Code 1   *Figure 2*. QR Code 2   *Figure 3*. QR Code 3

Mobile device. Apple iPad tablets running iOS 5 or later with built-in cameras (generation 2 or 3) were used for the treatment groups with AR (high AR and low AR). A custom software application was installed on each tablet prior to running the study to permit the participants to view the content in 3-D. Participants interacted with the 3-D images by aiming the tablet at each QR code. The tablets were wirelessly connected to the Internet so that corresponding questions could be queried from a remote database. Also, a wireless solution permitted multiple participants to launch the mobile AR learning activity and move around the classroom at the same time without the need for cables.

Learning activity. Depending on the participants' assigned treatment group, participants viewed and interacted with an arbitrary sample of fish within the designated

area by either looking around the room through a mobile device or looking at traditional paper-based methods of instruction. Images of the two-dimensional (2-D) paper-based images are shown in Figures 4, 5, and 6.

Figure 4. Non-AR Image 1   *Figure 5*. Non-AR Image 2   *Figure 6*. Non-AR Image 3

In either case, the content in the learning activity was exactly the same between all treatment groups (AR or no AR). The goal for each treatment group was also exactly the same; only the delivery method varied as described in the following section. The learning activity took on average 30 minutes to complete. It was designed to help learners understand proportions and sampling variability given a particular learning narrative. The learning activity covered statistical topics such as populations, samples, and what it means for a sample to be representative of the population. By analyzing fish of different color, participants analyzed varying samples and identified factors affecting sampling variability. The overall learning goal of the scenarios was to increase the statistical reasoning skills of the participants. Specifically, the learning objectives for the activity were:

- Build and describe sample distributions
- Calculate a probability statistic of a sample
- Recognize the variability between samples

48

- Describe the effect of sample size on how well a sample resembles a population

- Describe the differences between sample and population distributions

Learning narrative. The narrative for the learning activity placed the participants on a team of environmental scientists chosen to investigate a recent oil spill that happened near a small bayside town. Provided with a briefing on the behavior and treatment plan for the different types of oil, the participants were tasked with trying to pinpoint the source of the contamination by taking multiple samples of the fish from different drinking water sources as shown in Figure 7. As part of the narrative, participants were also provided instructions on how to complete the learning activity.
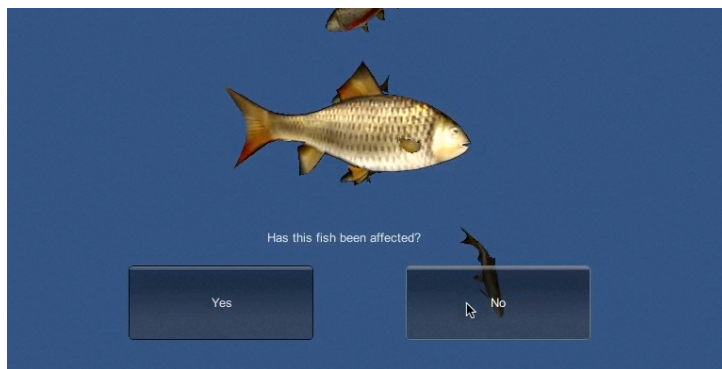
Figure 7. Screenshot of the selection of a fish with question prompt.

Treatment Conditions

High AR with collaboration. In this treatment group, participants were randomly assigned a partner to work with to complete the learning activity. As such, the pair shared a mobile device and navigated through the environment together. Unique to this treatment group, participants explored the physical space of the classroom to complete the learning activity by viewing and interacting with the QR codes that were located

equidistantly around the classroom as shown in Figure 8. These participants were allowed to confer to derive an answer for a given question during the learning activity; however, each participant was required to submit his or her responses to the pre- and posttest individually. For the learning activity, images were in the form of 3-D dynamic images and animations made with hand-rendered art and computer-generated objects. The learning activity was interactive and included other instructional text and traditional Apple iPad gesture inputs.

Figure 8. Participants working in pairs in the high AR condition.

High AR without collaboration. In this treatment condition, participants worked alone and navigated through the environment individually to complete the AR scenario as shown in Figure 9. With the use of a mobile device, participants had the ability to navigate around the room, but they were not permitted to discuss the scenario or exchange answers with other participants in the classroom. For the learning activity, the images were in the form of 3-D dynamic images and animations made with hand-rendered art and computer-generated objects. The learning activity was interactive and included other instructional text and gesture inputs.

Figure 9. Participants working without collaboration in the high AR condition.

Low AR with collaboration. In this treatment group, participants were randomly assigned a partner to work with to complete the learning activity. However, while sharing a mobile device, each pair was positioned at one stationary location to complete the learning activity as shown in Figure 10. These participants could confer to derive an answer for a given question during the learning activity; however, each participant was required to submit his or her pre- and posttest responses individually. For the learning activity, the images were in the form of 3-D dynamic images and animations made with hand-rendered art and computer-generated objects. The learning activity was interactive and included other instructional text and Apple iPad gesture inputs.

Figure 10. Participants working in pairs in the low AR condition.

Low AR without collaboration. In this treatment condition, participants worked alone to complete the learning activity while positioned at one stationary location. They were not permitted to discuss the scenario with others, nor exchange answers with other participants in the classroom. In the low AR scenario, participants had the same interactive AR experience with instructional text and Apple iPad gesture inputs.

Control group with collaboration. Participants were assigned to the control group to complete the exact same learning activity in pairs. However, to represent how the content is traditionally taught, the learning activity was done with pen and paper. All of the information was presented as traditional 2-D, static images, and the participants completed the activity at one stationary location. These participants could confer to derive an answer for a given question during the learning activity; however, each participant was required to submit his or her pre- and posttest responses individually. Participants assigned to this group were not permitted to utilize the application on the mobile device. All note taking and answers were recorded via pencil and paper, which were provided.

Control group without collaboration. Participants assigned to this control group completed the exact same learning activity with pen and paper; however, they worked alone to complete the learning activity. All of the information was presented as traditional 2-D, static images, and the participants completed the activity at one stationary location. Participants assigned to this group were not permitted to utilize the application on the mobile device. All note taking and answers were recorded via pencil and paper, which were provided.

Instruments

   Instruments for this study included a presurvey, pretest, posttest, and a perception survey that was administered after the learning activity was completed. These instruments were delivered using an online survey administration tool, Qualtrics, which is an application that allows for the creation of questionnaires and compilation of data. Participants completed the surveys individually at a designated workstation in a computer laboratory environment.

   Presurvey. Participants completed an online questionnaire (see Appendix A), which included the following elements:

- Demographics: gender, age, ethnicity, grade, and major

- Familiarity to learning task: how many statistics classes the participant has taken since the ninth grade (i.e., freshman statistics I, etc.)

- Statistics ability: ability to solve statistics problems related to probability and variance

- Attitude towards statistics: self-rating of positive, negative, or neutral

- Mobile device prior usage: how familiar the participant is with using an Apple iPad or other tablets

- Mobile device ability: self-rating of novice, intermediate, or expert tablet user

   Pretest and posttest. A brief online survey with 15 items was administered electronically to assess each participant's prior knowledge of the content area, statistical probability, sampling, and variability. This assessment was adapted from an activity by Rossman and Chance (2010) based on the curriculum from Glenberg, Glenberg, and Andrzejewski (2007). Directly before and after the learning activity, participants were

53

given the same 15-item test to measure their statistics knowledge beforehand and to assess what they learned about statistics from the learning activity. All of the items in the posttest were mapped to the learning objectives of the activity. Differences in scores from pretest to posttest were analyzed to evaluate if learning occurred. The researcher graded the assessment immediately after each session and provided the results to the participants via email after the data collection phase of this study, if requested. See Appendix B for the pretest and posttest questions and Appendix C for the posttest only questions.

Perception survey. To evaluate the participants' perceptions of collaboration and engagement, two scales were used for this study. To measure perceptions of collaboration during the learning activity, participants who were assigned to a treatment group with collaboration answered nine Likert-type questions created by the researcher related to communication, sharing, participation, and cooperation. See Appendix D for questions included in this survey. To measure perceptions of engagement during the learning activity, all participants answered 33 Likert-type questions related to focused attention, usability, aesthetics, endurability, novelty, and involvement. This scale was modified from one developed by O'Brien and Toms (2010). See Appendix E for questions included in this survey.

Cronbach's alpha was used to determine the internal consistency estimate for the reliability of the measures of collaboration and engagement. For the measure of perception of collaboration, the overall reliability of the scale was 0.83, which suggests reasonable reliability for the survey. For the measure of perception of engagement, the reliability of the scale was 0.91, which indicates good reliability for this survey.

Procedure

Students were invited to participate in this study through an announcement delivered via email for online courses or by an in-person announcement for face-to-face classes. An electronic announcement was made through a list of email addresses sourced from a participant pool managed by the administration in the education department at the university, called Sona System. The Sona System is a web-based human subject pool management software used for administering and organizing research studies. In-person announcements were made to various classes around the universities by obtaining instructor permission to recruit students who fit the demographics for participation in the study.

Once registered for the study, students selected a time from a list of available time slots based on the researcher's scheduling. On a rotating basis, treatment groups were scheduled by the day. For instance, for a given Monday session of participants, treatment Group 1 (high AR with collaboration) was run. On the following day, the next treatment group was run. This sequence was used until enough data for each treatment group was collected. In rare instances, full, intact classes of participants were run through the study. In those cases, the researcher would randomly assign participants to one of the six treatment groups by dividing up the number of participants equally. The breakdown of sample sizes per treatment group is shown in Table 2.

Table 2

Sample Size by Condition

	AR Experience	

	High AR	Low AR	No AR
Pairs	39	53	41
No Pairs	41	41	37

Participants were scheduled to attend a 30-minute session according to their assigned group to complete all of the activities of the study concurrently with no more than 10 participants assigned to the same treatment group. At the beginning of each session, participants were randomly assigned to one of the six treatment groups prior to implementing the treatments. Then, participants were given instructions for participating in the study. Before any data was collected, all participants were required to sign an electronic consent form agreeing to participate in the study. After completing the study, the researcher compensated each participant with either 10 extra credit points in a class or $10.00 in cash. The sequence of events for each session included the following procedures and elements:

1. Presurvey: After signing in and completing the consent form, participants completed the presurvey.

2. Pretest: Participants completed the pretest that:

    a. assessed the participant's incoming statistical knowledge, and

    b. measured how much learning took place as a result of the learning activity when considered in combination with the posttest.

3. Learning Activity Training: Participants were given the narrative for the learning activity and the learning activity instructions.

4.   Learning Activity: Participants completed the learning activity in accordance with his or her assigned treatment group.

5.   Posttest: After completing the learning activity, participants completed a posttest that:

   a.   assessed the participant's statistical knowledge, and

   b.   measured how much learning took place as a result from the learning activity when considered in combination with the pretest.

6.   Postsurvey: Participants completed the online surveys, which measured the participants' perceptions about collaboration and engagement.

7.   Debriefing and Checkout: Participants were debriefed on the study and received their compensation.

Chapter 3

Results

Primary Analysis by Variable

The results from three dependent variables were analyzed (learning gains, perception of collaboration, and perception of engagement). Table 3 presents a summary of each research question and analytic approach. For all statistical comparisons, the family-wise Type I error rate was set at the 0.05 level. Cohen's f was used as an effect size index, where 0.10, 0.25 and 0.40 correspond to small, medium, and large effect sizes, respectively (Cohen, 1988). Based on performance on the pretest, participants were categorized into three equally sized groups based on ability: high, average, or low prior knowledge. Each group was defined by their pretest scores. The high prior knowledge group was defined included those with pretest scores one standard deviation above the overall mean; the average prior knowledge group included those with pretest scores between one standard deviation above and one standard deviation below the overall mean; the low prior knowledge group included those with pretest scores one standard deviation below the overall mean. The group score ranges were 13 to 15 (high), 11 to 12 (average), and 1 to 10 (low).

Table 3

Research Questions and Analytic Approaches

Research Question	Data Source	Analyses
1. Does an AR experience combined with collaboration impact learning gains in participants' understanding of statistical reasoning concepts beyond a no	Pretest Scores Posttest Scores	Analysis of Variance

58

AR experience? Specifically: a. Does an AR experience (high AR, low AR, or no AR) produce differences in learning gains? b. Does working individually versus collaboratively produce differences in learning gains? c. Do different levels of AR experience and prior knowledge interact to produce differences in learning gains?		
2. Does a high or low AR experience facilitate the perception of a more collaborative experience than a learning experience without AR?	Collaboration Scale	Analysis of Variance
3. Does a high or low AR experience facilitate the perception of a more engaging experience than a learning experience without AR?	Engagement Scale	Analysis of Variance

Note. AR = augmented reality.

Prior Knowledge

A two-way analysis of variance (ANOVA) was conducted to evaluate whether participants' prior knowledge differed across the six treatment groups before participating in the study. The levels for the first factor, augmented reality (AR experience), were high AR, low AR, and no AR, while the levels of the second factor, collaboration, were pairs and no pairs. The dependent variable was pretest scores, and the means and standard deviations for the scores are presented in Table 4. The results indicated no significant main effects for AR experience $F(2,246) = 0.59$, $p = 0.55$, or collaboration $F(1,246) = 0.02$, $p = 0.88$, and there was no significant interaction, $F(2,246) = 0.42$, $p = 0.66$. Overall, the two-way ANOVA did not indicate statistical significance between the pretest

score means for each condition. Thus, the participants' prior knowledge did not differ across the six treatment groups prior to participating in the study.

Table 4

Mean Scores and Standard Deviation of Pretest Score

Collaboration	AR Experience	N	Mean*	SD
	High AR	36	11.17	2.30
Pairs	Low AR	53	11.53	2.32
	No AR	40	11.60	2.38
	High AR	44	11.39	2.20
No Pairs	Low AR	41	11.78	1.98
	No AR	38	11.26	2.55

Note. N = number of participants; * = Participant pretest mean scores out of 15; SD = standard deviation.

Learning Gains

A 3 x 2 ANOVA test was conducted to evaluate the effects of varying levels of AR experience and collaboration on learning gains of participants. The levels for the first factor, AR experience, were high AR, low AR, and no AR, while the levels of the second factor, collaboration, were pairs and no pairs. The dependent variable was learning gains, as measured by the difference in pretest and posttest scores. The means and standard deviations for the scores are presented in Table 5. The results for the 3 x 2 ANOVA indicated no significant main effects on learning gains from AR experience $F(2,246) = 2.22, p = 0.11, f = 0.02$, or collaboration $F(1,246) = 0.23, p = 0.63, f = 0.01$. Additionally, there was no significant interaction between AR experience and collaboration, $F(2, 246) = 0.53, p = 0.59, f = 0.01$. Overall, the 3 x 2 ANOVA did not indicate statistical significance between the difference score means of each condition.

Table 5

Collaboration and AR Experience Learning Gains Mean Scores and Standard Deviation

Collaboration	AR Experience	*N*	Mean	*SD*
	High AR	36	1.44	1.90
Pairs	Low AR	53	0.98	1.79
	No AR	40	0.58	1.53
	High AR	44	0.98	2.41
No Pairs	Low AR	41	1.07	1.78
	No AR	38	0.61	1.85

Note. N = number of participants; SD = standard deviation.

A two-way ANOVA test was conducted to evaluate the effects of varying levels of AR experience with regard to the moderating variable—prior knowledge—on learning gains of participants. The levels for the first factor, AR experience, were high AR, low AR, and no AR, while the levels of the second factor, prior knowledge, were high, average, and low. The dependent variable was learning gains, as measured by the difference in pretest and posttest scores. The means and standard deviations for the scores are presented in Table 6. The results for the two-way ANOVA indicated significant main effects on learning gains from AR experience $F(2,243) = 3.27$, $p = 0.04$, $f = 0.03$ and prior knowledge $F(2,243) = 34.37$, $p < 0.001$, $f = 0.22$. Additionally, there was a significant interaction between AR experience and prior knowledge, $F(4, 243) = 2.84$, $p = 0.025$, $f = 0.05$.

Table 6

AR Experience and Prior Knowledge Mean Scores and Standard Deviation of Learning Gains

AR Experience	Prior Knowledge	*N*	Mean	*SD*
High AR	High	26	-0.12	1.40
	Average	30	0.93	2.00
	Low	24	2.92	2.06
Low AR	High	37	0.03	1.28
	Average	30	0.97	1.56
	Low	27	2.44	1.67
No AR	High	29	0.10	0.94
	Average	24	0.67	1.40
	Low	25	1.08	2.38

Note. N = number of participants; SD = standard deviation.

Because the interaction between AR experience and prior knowledge was significant, the main effects were ignored; instead, the simple main effects were examined. Thus, the differences between prior knowledge across the three types of AR experience (high AR, low AR, and no AR) were examined separately. To control for a Type I error for each of the three pairwise comparisons, the modified Shaffer (1986) method was used with the alpha level set at 0.0167 (0.05/3) a priori. The first comparison tested the hypothesis that the difference in mean learning gains in the high AR condition was the same across the three levels of prior knowledge. The comparison was significant, $F(2, 243) = 21.34$, $p < 0.001$, $f = 0.15$. The second comparison tested the hypothesis that the difference in mean learning gains in the low AR condition was the same across the three levels of prior knowledge. The comparison was significant, $F(2, 243) = 16.55$, $p <$

0.001, $f = 0.12$. The third comparison tested the hypothesis that the difference in mean learning gains in the high AR condition was the same across the three levels of prior knowledge. The comparison was not significant, $F(2, 243) = 2.36$, $p = 0.10$, $f = 0.02$. The results of these comparisons suggests that participants with low prior knowledge had higher learning gains in the high AR and low AR conditions than do participants with high or average prior knowledge.

Participant Perceptions

Collaboration. A one-way ANOVA test was conducted to evaluate the effects of varying levels of AR on participants' perception of collaboration. The levels for the factor were high AR, low AR, and no AR, and the dependent variable was collaboration score. The means and standard deviations for the perception of collaboration scores are presented in Table 7. The results for the one-way ANOVA did not indicate a significant main effect from AR experience $F(2,124) = 0.76$, $p = 0.47$, $f = 0.01$ on perception of collaboration scores.

Table 7

Perception of Collaboration Mean Scores and Standard Deviation of Scores

AR Experience	N	Mean*	SD
High AR	38	7.69	1.39
Low AR	49	7.96	1.60
No AR	40	7.53	1.92

Notes. N = number; SD =standard deviation; *Participant self-reported mean scores reported on a scale of (1 = Poor, 10 = Excellent).

Engagement. A one-way ANOVA test was conducted to evaluate the effects of varying levels of AR on participants' perception of engagement. The levels for the factor

were high AR, low AR, and no AR, and the dependent variable was engagement score. The means and standard deviations for the perception of engagement scores are presented in Table 8. The results for the one-way ANOVA indicated significant main effects from AR experience $F(2,246) = 10.84, p = 0.001, f = 0.08$ on perception of engagement scores.

Follow-up tests were conducted to evaluate pairwise differences between the means across the levels of AR experience. To control for Type I error for each of the three simple main effects, the modified Shaffer (1986) method was used with the alpha level set at 0.0167 (0.05/3) a priori. There was a significant difference in means between the high and no AR experience conditions, and low and no AR experience conditions, but no significant differences between the high and low AR experience conditions. Participants assigned to either high or low AR experience conditions reported higher perception of engagement scores than those who were assigned to the no AR experience conditions.

Table 8

Perception of Engagement Mean Scores and Standard Deviation of Scores

AR Experience	N	Mean*	SD
High AR	79	7.49	1.39
Low AR	92	7.58	1.22
No AR	78	6.67	1.55

Notes. N = number; SD =standard deviation; *Participant self-reported mean scores reported on a scale of (1 = Poor, 10 = Excellent)

Chapter 4

Discussion and Conclusion

One of the main goals of this study was to investigate the impact of different

levels of augmented reality (AR) experience and collaboration on students' learning and

their perception of engagement and collaboration in statistics education. The following

section reviews the findings relative to the study's research questions.

Findings by Research Question

Research Question 1. Does an AR experience combined with collaboration

impact learning gains in participants' understanding of statistical reasoning concepts

above a no AR experience?

The impact of AR experience was examined, namely a high AR experience

(interacting with three-dimensional, or 3-D, images coupled with movement through a

physical space), a low AR experience (interacting with 3-D images without movement),

and no AR experience (two-dimensional, or 2-D, images and text without movement), as

well as two levels of collaboration (pairs and no pairs). The purpose of this analysis was

to investigate whether or not an AR experience, combined with collaboration, would

provide a more meaningful context to increase learning gains. The differences in the

mean scores from the pretest to the posttest were used to assess participants' learning

gains to evaluate the effects of the AR experience. The anticipated outcome was that

participants who completed the learning activity in the high AR experience with

collaboration would have higher learning gains than participants in the other conditions.

This suggested that a participant's learning gains in the high AR experience with

collaboration did not differ from the other conditions. In fact, there were no significant

65

differences between any of the AR experience conditions taking into account collaboration. Previous research on AR and collaboration has already showed signs of promise when combined to support learning (Kaufmann & Schmalstieg, 2003; Yuen, Yaoyuneyong, & Johnson, 2011). However, the related analysis in this study did not support this notion and did not return significant results. Despite having data from as many as 252 participants to attempt to effectively analyze the effects of AR experience and collaboration together in this study as initially intended, the six experimental conditions perhaps reduced the opportunity to find statistical significance in the study. Based on the small reported effect sizes, dividing the number of participants per condition potentially precluded the analysis from finding significance given the number of comparisons. The effect size associated with this analysis was minimal for AR experience ($f = 0.02$) and collaboration ($f = 0.01$), which indicates that participant learning gains were equivalent to one another across the six conditions. However, effect sizes other than small, by Cohen's (1988) classification, are difficult to achieve during educational interventions according to authors Lipsey and Wilson (1993). This is especially the case for evaluating the impacts on learning gains (Lipsey & Wilson, 1993). The combination of AR experience and collaborative learning through the use of AR to increase learning gains in education may be still worthwhile. However, the results from this study were inconclusive. The recommendation would be to pursue alternative ways to investigate AR experience with collaboration in a future study.

Research Question 1a. Does an AR experience (high AR, low AR, or no AR) produce differences in learning gains?

The impact of the levels of AR experience was examined, namely a high AR experience, as well as three levels of prior knowledge (high, average, and low). The purpose of this analysis was to investigate whether or not an AR experience would provide a more meaningful context to increase learning gains. Two outcomes were anticipated: participants who completed the learning activity in either the high or low AR experience would have higher learning gains than participants who completed the learning activity without an AR experience, and participants who completed the learning activity in the high AR experience would have higher learning gains than participants in the low AR experience.

Neither outcome was supported by the results. The 3 x 2 analysis of variance (ANOVA) indicated that there was not a significant difference in learning gains across the three types of AR experiences. This could be attributed to the strong role the moderating variable, prior knowledge, had on AR experience. Design research in the related area of simulations and games suggests that learning gains vary according to prior knowledge (Bower, Kelsey, & Moretti, 2011; Mayer, 1997; Park, Lee, & Kim, 2009; Plass, et al., 2011). Studies focused on AR have also shown evidence that augmented learning experiences are potentially dependent on level of prior knowledge (Elinich, 2011; Klopfer & Squire, 2008; Shelton, 2003). This could help explain the moderating effect of prior knowledge on AR experience in this study, which is supported later in this section. The interaction effect between AR experience and prior knowledge in comparing learning gains is consistent with the interpretation that prior knowledge could serve a more important role in determining learning gains from AR experiences.

Research Question 1b. Does working individually versus collaboratively produce differences in learning gains?

The impact of the two levels of collaboration (pairs and no pairs) was examined. The purpose of this analysis was to investigate whether or not participants working in pairs would increase learning gains. It was the expectation that because of the form factor of the mobile device and the interactivity of the AR experience (high and low), participants would feel that they had more opportunity to collaborate, specifically for sharing the device, communicating, participating, and cooperating during the learning activity.

The results from this examination suggested that participants' learning gains were not different depending on which condition they were in: pairs and no pairs. Though working in pairs did not have a significant impact on learning gains, the value of collaboration should not be dismissed. The difference between the mean scores of those who completed the learning activity in pairs as compared to those who worked alone improved marginally; regardless, it was statistically insignificant. The effect size associated with this analysis was minimal ($f = 0.01$), which indicates that participant learning gains were equivalent to one another. The vast body of literature suggests that collaboration positively impacts learning more so than working individually. However, very few studies have been able to empirically show significant differences on learning gains (Chen & Zhang, 2013). Within the scope of this study, collaboration did not directly impact learning gains; however, the effects may have been dependent upon the extent to which pairs actually engaged in productive interactions during the learning activity (Dillenbourg, Järvelä, & Fischer, 2009). Therefore, further examination with

collaboration as the primary variable is suggested to be able to substantiate the impact on participants' learning gains.

Research Question 1c. Do different levels of AR experience and prior knowledge interact to produce differences in learning gains?

The impact of AR experience across the three levels of prior knowledge (high, average, and low) was examined, as well as the impact of prior knowledge across the three types of AR experience (high AR, low AR, and no AR). The purpose of this analysis was to investigate whether or not learning gains in each type of AR experience was dependent on prior knowledge. The anticipated outcome was that participants with low prior knowledge would show higher learning gains than participants with high or average prior knowledge.

The results indicated that an interaction did exist between AR experience and prior knowledge. This suggests that the impact of AR experience did depend on the level of prior knowledge. Further investigation into this interaction indicated that participants with low prior knowledge had different learning outcomes depending on which condition they were in: high AR, low AR, and no AR. Participants categorized as having low prior knowledge benefited the most from a high AR experience. The same benefit was not apparent between those categorized as high and average prior knowledge. A possible explanation for this occurrence is that participants with low prior knowledge had the most room to grow in this content area, thus having the highest learning gains overall. Because prior knowledge is considered one of most important predictors of future learning (Beier & Ackerman, 2005; Dochy, Segers, & Buehl, 1999), it is understandable that this moderating variable was responsible for the interaction between learning gains and AR

69

experience. Similarly, Schwartz, Sears, and Chang (2007) made the case to consider

students' prior knowledge when conceiving ways to help motivate and engage students to

develop statistical reasoning skills. The previous research results from the statistics

education domain have produced mixed results about the sensitivity novice and advanced

learners have to new pedagogical and technological interventions.

However, that does not fully explain why increased learning gains for those with

low prior knowledge occurred between the high AR and no AR conditions, and low AR

and no AR conditions. One potential explanation for this finding is, based on the theory

of embodiment, that participants in the high and low AR experience conditions used more

of their physiological senses to form deeper cognitive connections with the information

for later recall. Embodiment as a learning theory explains how our understanding of the

world around us is formed through our bodies, not just our sight and hearing (Lakoff &

Johnson, 1980). Because participants' learning gains with low prior knowledge in the

high AR experience were not significantly different from those in the low AR experience,

it could also have been the case that the movement around a physical space alone was not

enough to produce significant learning gains between the high and low AR experiences.

Perhaps just the overlap of virtual objects on the real-world transformed what are

considered abstract ideas into concrete metaphors for the participants to help facilitate

new knowledge acquisition, or schemas. AR experiences as a whole, with or without

movement, could potentially help develop new schemas to encode the structure and

relationships experience during the learning activity (Xu et al., 2011). In context, these

schemas could then be adapted and applied to future experiences through the use of the

metaphor (Lakoff & Johnson, 1980). Recent developments in the research related to

70

embodiment also suggest that, unlike other technologies, AR can potentially increase learning gains through the extra inclusion of multiple senses to anchor new information cognitively (Bujak et al., 2013; Cuendet, Bonnard, Do-Lenh, & Dillenbourg, 2013; Holz et al., 2011; Jetter, Geyer, Schwarz, & Reiterer, 2012). Situated in a particular environment (i.e., a classroom), AR experiences show potential to provide embodied learning experiences that are natural and intuitive where the body and senses are used to interacting with objects within the learning environment (Holz et al., 2011; Xu et al., 2011).

Another reasonable explanation for the interaction could be that the inclusion of movement around a physical space was important for the cognitive process for those with low prior knowledge more so than for others with high or average prior knowledge. Supported by the embodied cognition literature, movement should be incorporated into the AR experience if it is for students with low prior knowledge in the content area (Johnson-Glenberg, Birchfield, Tolentino, & Koziupa, 2013). In terms of the pedagogy, an embodied experience can be provided by the use of situated context. A significant finding from this study was that participants who were categorized as having low prior knowledge showed the largest learning gains across the three types of AR experiences. This benefit was particularly pronounced between the high AR treatment condition and the no AR treatment condition where the results indicated a statistically significant difference. In terms of technology, an embodied experience can be provided by the use of mobile devices with gesture computing and wireless connectivity (Jetter et al., 2012). Under these circumstances, a high or low AR experience like those used in this study

could be a beneficial tool for teaching basic statistical concepts in a university level introductory statistics course.

Research Question 2. Does a high or low AR experience facilitate the perception of a more collaborative experience than a learning experience without AR?

To evaluate the impact of the perception of collaboration using tablets, a subsample of participants (127 out of 252) was randomly selected to participate in the study as pairs. Collaboration perceptions data was collected from these individuals in the subsample using the customized collaboration scale, but such data was not collected from the other participants. The expectation was that due to the form factor of the mobile device plus the movement about the space, participants would feel that they had more opportunities to collaborate.

The results were inconclusive. The scores from the collaboration scale did not indicate a difference between participants' perceptions in the high or low AR experience as compared to those with no AR. There was no difference in participants' perceptions of collaboration between the conditions of this study. Similar to the conclusions of the previous analysis, this is not to say that collaboration should not be considered when designing learning activities for statistics education. Instead, the suggestion for future researchers is to consider the many other aspects of collaborative learning theory, such as the number of group members, roles and responsibilities, and the types of communication when building a learning activity for statistics education.

Research Question 3. Does a high or low AR experience facilitate the perception of a more engaging experience than a learning experience without AR?

The study examined the impact of an AR experience on perceptions of engagement. The purpose of this analysis was to investigate whether or not the contextualized activity and the interaction with the technology of the AR experience (high and low) would result in the participants perceiving higher engagement, as measured by the adapted engagement scale (O'Brien & Toms, 2010). It was predicted that participants who completed the learning activity in an AR experience (high or low) would report a higher perception of engagement than participants who complete the learning activity without any AR.

The results from this analysis supported the hypothesis. Participants that completed the learning activity in an AR experience reported a higher perception of engagement as measured by the engagement scale. The connection between AR and engagement could have been the result of the newness of the experience. Overall, the majority of participants had little or no experience working with tablets. This could have resulted in participants being more engaged due to the novelty of the technology. On the other hand, this result could stem from the interactivity and visual richness generated by a high or low AR experience. Interactivity and visual elements are some of the benefits innate to AR that could help explain the increased perception of engagement reported in this study. Engagement is considered important due to its instrumental role in improving learning gains (Dunleavy, Dede, & Mitchell, 2009).

Limitations

There were a few limitations within the study worth noting. The learning activity that served as the foundation of this study revolved around understanding basic statistical concepts including probability, sampling, and variability. Although the content is heavily

73

connected to the literature, it could be the case that the learning objectives for the level of students in this study missed the mark. Having not prescreened the participants prior to participating in the study, it is difficult to determine if they were over- or underprepared for the content. Moreover, the learning activity itself turned out to take only approximately 20 minutes to complete instead of the initial 30 minutes allotted. Therefore, it is difficult to state with confidence that the learning activity was sufficient to produce meaningful learning gains.

Another potential limitation of the study was the design of the materials accompanying the learning activity. As described in the Method section, to assess statistical reasoning ability, the participants were charged with investigating a fictitious oil spill by sampling different areas of a bay. To be consistent, the instructions were completely text-based. At each area, participants were supposed to read the text-based placards placed next to the targets that gave them clues to help them answer questions presented later on the posttest. On more than a few occasions, participants showed signs of not seeing or paying attention to the placards. This was especially true of the participants assigned to the AR treatment groups. In general, once the participants picked up the tablet, their attention was focused mostly on the AR application. Also, those assigned to the no AR treatment group had the tendency to simply focus on the images of the fish, not the placards. A better design and placement of the instructional materials would have been to put the placards closer to the activities in a way that increased the chances of the participants seeing them. In hindsight, different images of the areas of the bay may have made a difference. There is a possibility that some of the participants had a conflicting interpretation of what the image represented and where they were sampling.

For instance, when the participants were sampling the middle part of the bay, the image could have been visually construed as a small lake. This issue potentially affected the ability to assess the participants' ability to statistically reason when the participants were asked questions related to the overall task of investigating the oil spill.

As previously described, half of the participants were assigned to complete the learning task in pairs. These assignments were completely randomized to alleviate the possibility of two friends working together, a pairing which would have potentially skewed the perception of collaboration data. However, this possibly affected the outcome of the study since, in some cases, one member of the pair finished the presurvey faster than the other. As a result, one participant had to wait sometimes as long as five to seven minutes before starting the learning task. It is possible that this led to some of the participants forgetting the instructional material. Participants were assigned to pairs before the pretest to avoid the possibility of partnering two participants together who completed the pretest quickly. It could be that those who finished first were more familiar and comfortable with the content, which could also bias the results of the study.

Also, the data collected through the self-report questionnaires could be viewed as a limitation of the study. Due to the number of participants included in the study, it was difficult to consider researcher observations or other scientific methods of collecting data related to engagement and collaboration. Instead, participants provided information based on their own perceptions. While there is no evidence that participants attempted to provide misleading information, it is difficult to discern if the information provided by participants was factual or free from honest mistakes. In future studies, it would be worthwhile to consider how alternative or mixed-methods approaches might support data

validity. For instance, qualitative and biometric data could provide supporting information beyond the self-report measures used in this study.

General Discussion

This study uncovered a strong relationship between AR experience and prior knowledge for teaching basic statistical concepts. The strong relationship suggests that it is important to consider how visualizations are used in the process of learning and how different visual representations are utilized for students who have varying levels of prior knowledge and possess different learning attributes (Shelton, 2003). Furthermore, prior knowledge should continually be a consideration for any learning experience where students are tasked with constructing new knowledge or understanding tied to what they already know (Elinich, 2011). In this study, students constructed their own knowledge by testing existing ideas and approaches based on their prior knowledge and experiences, and actively applied these to a new situation. They also constructively integrated the new knowledge gained with preexisting intellectual constructs (Ryu & Parsons, 2009). The active participation of the learning activity facilitated the inclusion of the prior knowledge, attitudes, habits, and interests that the students brought to the experience. Students who participate in augmented learning experiences can draw on the framework of the activity and their own knowledge and imagination as they experience it. They have the ability to act and respond as though the learning activity is real, even if there is very little explicit visual support for the metaphor of it (Colella, 2009).

Lastly, the aim of the AR experience designed for this study was to help students better understand basic statistical concepts in the form of a metaphor by superimposing 3-D images onto real-world objects in a meaningful way. This study helps extend the

potential of this technology to statistics education. As posited in the review of the literature, AR as a whole and the supporting technology (i.e., tablets, wireless connectivity, etc.) demonstrated enough stability to be implemented for projects of this scale. Moreover, this study promisingly indicated that if given the technology and a specific context, a positive impact could be made on learners' learning gains and level of engagement.

The results of this study also documented a relationship between learning gains and prior knowledge, but found no such relationship between learning gains and collaboration. The AR learning activity positively impacted learning gains for participants with lower prior knowledge, and AR also positively impacted participants' perceptions of engagement. Thus, it would be strategic to continue to focus on the relationship between learning gains and engagement for research ideas involving the design and implementation of new AR systems. Additionally, further investigation is necessary to adequately address the relationship between learning gains, embodiment, and collaboration.

Based on these results, a new approach to statistics education at the college level is seemingly needed. An approach that is specifically suited for students with low prior knowledge will help them develop a deeper understanding of statistical reasoning. "Rather than present material in a linear fashion, as most textbooks and current courses do," a new framework is needed to help create fresh and innovative approaches to teaching sampling, probability, and variability (Garfield delMas, & Chance, 2007, p. 121).

Future Research Directions

As the demand increases for new and innovative ways to teach complex, higher order cognitive skills, it seems appropriate to conduct research for designing and implementing AR technology in educational settings. In addition to the amalgamation of already established research, the work presented in this study shows an opening into the education arena of developing possibilities for AR. However, AR is a bourgeoning area of research for the educational field, and there are still more questions to be answered. Is the educational application of the AR experience designed and implemented for this study more effective than traditional instructional methods to the degree that it is ready to be implemented in the classroom? There are encouraging signs, yet much more empirical evidence is required before that question can be definitively answered. In the context of this study, educational applications of AR should focus on how AR could be used for learning, how students learn from it, and what they actually learn.

Impact of AR on spatial cognition. Other possibilities for studies that would extend the body of research on AR in education include identifying new ways to create embodied learning experiences while leveraging the natural affordances of AR and mobile technology. As described, one of the limitations of this study was the inability to genuinely identify the impact of having users move about a physical space while interacting with virtual objects, which allows users to take advantage of the fusion of human senses that could be incorporated for spatial cognition.

A number of difficulties arose concerning statistical and methodological issues that hindered reaching any strong conclusion in regards to spatial cognition. It might be worthwhile to consider another study that takes into account the criticism raised in this study. An added suggestion would be to enunciate the users' space both physically and

virtually so the user has the association of a virtual location based on their physical location. From what is already known from the embodied cognition literature, the benefit(s) of effectively doing so could potentially have ramifications on learning gains for learners.

Collaborative AR learning experiences. Research studies that use AR in collaborative settings are evolving as more interesting mobile devices are available. New applications are being created to provide more engaging field trips or museum experiences. AR technology can be useful for this purpose by helping in the development of a learning environment that exhibits a meaningful context through the use of visualizations. However, most of the current research is aimed at the development of the systems themselves, rather than empirical work on how the systems have been used for educational purposes.

Situated AR learning experiences. In a few initiatives, researchers have reported results from studies that directly correlate with situated learning activities. Some feedback from those studies indicate the perception that AR experiences provide a cool factor in the way a learner can investigate real-world problems in a mixed-reality world. It would also be interesting to continue to explore the impact of the cool factor on engagement and learning. An example of a study would be assigning some participants to act in the role of agents or scientists and leave others without a role to see how these roles impact the results.

Educational application of AR in other content areas. Although research on augmented learning experiences constitutes an expanding research domain, claims that AR truly enhances student understanding of statistical concepts still lacks solid empirical

evidence. Despite the various studies conducted to date, there is clearly a need for more thorough evaluation of the motivational and learning effects of augmented learning experiences on developing higher order cognitive skills like statistical reasoning. Experimental studies in which AR is compared to traditional instructional methods for teaching statistical concepts are needed.

In terms of content areas, there is no current evidence to specify that the same or similar AR systems could not be used for other educational disciplines. The combination of real and virtual objects from AR could be used to recreate an archeological dig to uncover bones, just as a paleontologist would. It could also be used in the field of medicine to train future doctors by creating practice opportunities with interactive biological systems, which are difficult and expensive to see otherwise. Another idea for a follow-up study is to use AR for teaching concepts that involve dynamic, changing relationships, such as learning about covalent bonds in chemistry or supply and demand models in economics class. Instead of traditional passive lecturing methods, learners could actively be involved while exploring different cause-and-effect concepts. If the building of knowledge through real events in the world must be carried out through interaction with objects in that world, then learners must experiment with what is in their current environment. It is a necessary part of internalizing what is external in the world, and takes on meaning in the form of objectification, experience, and action (Price & Rogers, 2004).

Before any AR-based learning activity is rolled out at any level, the technology needs to be tested in teachers' hands. If teachers are not comfortable with the technology, there is a lower chance that the students will be successful using it. In future research, it

is interesting to consider an AR system that could be adapted by the instructor through a website that would allow them to customize and generate new learning experiences that are easily integrated into the instruction. The system could be designed to allow the teacher to change the questions asked and provide possible answer choices. It could also serve as a dashboard for the teacher to help assess the students' performance and understanding of intended concepts.

Reliability of instruments. To better measure student perceptions, the development and implementation of new and existing instruments should be tested. This study used self-reported data to evaluate participants' perceptions from AR experience and collaboration. Other empirical methods include observations and interviews; however, more work needs to be done to establish reliable instruments of measurement specifically for AR. So far, different instruments have been used that were originally designed for other technologies. It would be helpful to have instruments that are designed and applied across educational applications for AR.

AR as a performance support tool. Creating self-guiding performance support systems is another popular application of AR that is currently gaining attention. Performance support systems are designed to allow users the ability to access the information needed to complete his or her job at a moment's notice (Gal & Nachmias, 2011). Due to the embedded affordances of the technology, it is foreseeable that AR could potentially be a great performance support tool to provide interactive guidance that directs a student through a learning task. Instead of referring to static text and images, students could view and interact with worked examples to solve math or science problems. Also, a mobile AR system could be designed for the purposes of providing

real-time instructions while performing either a motor-skill activity like students

navigating between classes around a campus or for an engineering student following an

assembly schematic.

The medical profession has already pursued AR as a performance support tool.

There is an AR system called Augmented Intervention Assistant (AIA) that allows

physicians to practice diagnosing patients using a head-mounted AR display (Kreiser,

2012). The heads up display (HUD) is used in conjunction with a generic AR system that

basically consists of a mobile device with a camera that fits on the user's head. The

camera recognizes the symptoms and then displays related feedback on the user's display

screen. If the user moves the camera on the HUD over a patient and his or her symptoms,

the 3-D image on the screen will provide suggested actions to treat possible ailments.

Besides providing digital feedback in real-time, a future study might investigate how the

AIA, or a similar AR system, could potentially improve a physician's ability to diagnose

and treat patients. The addition of technological advances such as eyewear from Epson,

Vuzix, ReconJet, or GlassUp could help carry AR into the mainstream as a popular

computing accessory.

This concept of AR as a performance support tool could be extended to other

professions as well. An AR application could also enhance the guidance process to

employees in the workforce who regularly complete complex tasks such as engineers and

scientists. Receiving just-in-time feedback could be a more flexible performance support

solution than other training solutions (Gal & Nachmias, 2011). A study designed to

investigate how AR can offer real-time learning and performance support in the

workplace is worth considering. Research concerning how AR could be used as a

performance support tool to provide just-in-time help beyond just the classroom could help inform what we currently know.

Field testing AR in the classroom. Finally, and maybe most importantly, to be able to truly know the impact of any AR system built for educational or performance support purposes, it is recommended that studies be performed in authentic situations. For instance, trialing any of the notable AR systems from the literature in an actual classroom context over time would be important. Perhaps a longitudinal study would be ideal for looking at the many aspects that could help determine the true viability of the system in an actual classroom with actual students and teachers over a semester or academic year. Currently, most studies describe education applications of AR systems for succinct or one-time learning activities. These are important first steps of any exploration, but the examination of long-term effects are subsequently needed.

Conclusion

This study leveraged a constellation of pedagogical approaches including collaborative learning, embodied cognition, and situated learning to propose and examine a new framework for using AR to teach statistical reasoning with mobile devices. The research involved college students completing a learning activity that required the application of statistical reasoning skills to solve a contextualized problem. Through the exploration of new and innovative ways while allowing students to move about a physical space and interact with virtual and real objects, the study found encouraging results. Prior knowledge played more of an important part of this study than anticipated in determining the effects of an AR experience, whereas collaboration did not. After categorizing participants by prior knowledge, the results were clear. Participants with low

83

prior knowledge showed significant learning gains after participation in a high or low AR experience. Conversely, the results showed no statistical difference between students assigned to work in pairs versus those who worked individually. As an integral component of learning, students' perception of engagement was also positively impacted from either a high or low AR experience as measured by the O'Brien and Toms (2010) engagement scale. However, the differences between students' perception of collaboration in each condition was not statistically significant.

Embodied cognition, collaboration, and perceptions of collaboration and engagement were all comprehensively analyzed together to evaluate the impact of an augmented learning experience—an analysis unique to this study. It is also one of the first studies to look at both undergraduate and graduate students' statistical reasoning skills. This current study is one of the first to show empirical evidence that AR can be designed to support learning in the domain of statistics education. Students with low prior knowledge especially benefited from a high AR learning experience. The research for this study was clear in that low-knowledge learners were aided the most from a high AR experience with movement within the physical space.

The results from the study present new knowledge about AR and its use in an educational setting. Overall, the study suggests that AR directly and positively impacts learning gains while providing interactive, hands-on practice opportunities working with statistical data in a meaningful way for students with low prior knowledge. Further evaluation of the implications that AR experiences have on spatial cognition, situated learning, high order skill development, performance support, and classroom applications are needed and could build from this study.

REFERENCES

Anderson, J. R., Reder, L. M., & Simon, H. A. (1996). Situated learning and education. *Educational Researcher, 25*(4), 5–11. doi:10.3102/0013189X025004005

Anderson, M. L. (2003). Embodied cognition: A field guide. *Artificial Intelligence, 149*(1), 91–130. doi:10.1016/S0004-3702(03)00054-7

Azuma, R. T. (1997). A survey of augmented reality. *Presence: Teleoperators and Virtual Environments, 4*(6), 355–385.

Azuma, R., Bailot, Y., Behringer, R., Feiner, S., Julier, S., & Macintyre, B. (2001). Recent advances in augmented reality. *IEEE Computer Graphics and Applications, 21*(6), 34–47. doi:10.1109/38.963459

Barab, S., & Dede, C. (2007). Games and immersive participatory simulations for science education: An emerging type of curricula. *Journal of Science Education and Technology, 16*(1), 1–3. doi:10.1007/s10956-007-9043-9

Barab, S., Sadler, T. D., Heiselt, C., Hickey, D., & Zuiker, S. (2010). Erratum to: Relating narrative, inquiry, and inscriptions: Supporting consequential play. *Journal of Science Education and Technology, 19*(4), 387–407. doi:10.1007/s10956-010-9220-0

Barab, S., Thomas, M., Dodge, T., Carteaux, R., & Tuzun, H. (2005). Making learning fun: Quest Atlantis, a game without guns. *Educational Technology Research and Development, 53*(1), 86–107. doi:10.1007/BF02504859

Batanero, C., & Dıaz, C. (2011). Training school teachers to teach probability: Reflections and challenges. *Chilean Journal of Statistics, 3*, 1–15.

Beier, M. E., & Ackerman, P. L. (2005). Age, ability, and the role of prior knowledge on the acquisition of new domain knowledge: Promising results in a real-world learning environment. *Psychology and Aging, 20*(2), 341–355. doi:10.1037/0882-7974.20.2.341

Billinghurst, M., & Kato, H. (2002). Collaborative augmented reality. *Communications of the ACM, 45*(7), 64–70. doi:10.1145/514236.514265

Bujak, K. R., Radu, I., Catrambone, R., MacIntyre, B., Zheng, R., & Golubski, G. (2013). A psychological perspective on augmented reality in the mathematics classroom. *Computers & Education, 68*, 536–544. doi:10.1016/j.compedu.2013.02.017

Chan, M. S., & Black, J. B. (2006). Direct-manipulation animation: Incorporating the haptic channel in the learning process to support middle school students in science learning and mental model acquisition. In S. Barab, K. Hay, & D. Hickey (Eds.), *Proceedings of the 7th International Conference on Learning Sciences* (pp. 64–70). Bloomington, IN: International Society of the Learning Sciences.

Chance, B. L. (2002). Components of statistical thinking and implications for instruction and assessment. *Journal of Statistics Education, 10*(3), 1–14. Retrieved from www.amstat.org/publications/jse/v10n3/chance.html

Chance, B., Ben-Zvi, D., Garfield, J., & Medina, E. (2007). The role of technology in improving student learning of statistics. *Technology Innovations in Statistics Education Journal, 1*(1), 1–26.

Chance, B., Delmas, R., & Garfield, J. (2004). Reasoning about sampling distributions. In D. Ben-Zvi & J. B. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 295–323). Dordrecht, Netherlands: Kluwer Academic Publishers.

Chapman, P., Selvarajah, S., & Webster, J. (1999). Engagement in multimedia training systems. In *Proceedings of the 32nd Annual Hawaii International Conference Systems Sciences, 1999 (HICSS-32).* (pp. 1084–1093). Maui, HI: IEEE. doi:10.1109/HICSS.1999.772808

Chen, Q., & Zhang, J. (2013). Collaborative discovery learning based on computer simulation. In A. M. O'Donnell, C. E. Hmelo-Silver, & G. Erkens (Eds.), *Collaborative learning, reasoning, and technology* (pp. 127–146). New York, NY: Routledge.

Chick, H., & Watson, J. (2001). Data representation and interpretation by primary school students working in groups. *Mathematics Education Research Journal, 13*(2), 91–111. doi:10.1007/BF03217101

Christian, J. (2006). *Augmented reality in corporate pervasive e-education: Novel ways to support aviation maintenance training.* Paper presented at the Innovation North

Research Conference. Leeds, UK. Retrieved from
http://www.researchgate.net/publication/251795985

Cobb, G. W., & Moore, D. S. (1997). Mathematics, statistics, and teaching. *The American Mathematical Monthly, 104*(9), 801–823. Retrieved from http://www.jstor.org/stable/2975286

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. New York, NY: Routledge.

Colella, V. (2009). Participatory simulations: Building collaborative understanding through immersive dynamic modeling. *Journal of the Learning Sciences*, *9*(4), 471–500.

Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition, 106*(2), 1047–58. doi:10.1016/j.cognition.2007.04.010

Costabile, M. F., De Angeli, A., Lanzilotti, R., Ardito, C., Buono, P., & Pederson, T. (2008). Explore! Possibilities and challenges of mobile learning. In *Proceedings of the 26th Annual SIGCHI Conference on Human Factors in Computing Systems*. (pp. 145–154). Florence, Italy: ACM.

Cuendet, S., Bonnard, Q., Do-Lenh, S., & Dillenbourg, P. (2013). Designing augmented reality for the classroom. *Computers & Education, 68*, 557–569. doi:10.1016/j.compedu.2013.02.015

Dede, C. (2005). Planning for neomillennial learning styles: implications for investments in faculty and technology. In D. Oblinger & J. Oblinger (Eds), *Educating the Net generation* (pp. 15.1–15.22). Boulder, CO: EDUCAUSE. Retrieved from http://www.educause.edu/educatingthenetgen

Dede, C. (2009). Immersive interfaces for engagement and learning. *Science, 323*(66), 66–69. doi:10.1126/science.1167311

delMas, Robert, C., Garfield, J., & Chance, B. L. (1999). A model of classroom research in action: Developing simulation activities to improve students' statistical reasoning. *Journal of Statistics Education, 7*(3), 1–18.

Dillenbourg, P. (1999). What do you mean by collaborative learning? In P. Dillenbourg (Ed.), *Collaborative-learning: Cognitive and computational approaches*.

Advances in Learning and Instruction Series. (pp. 1–19). New York, NY: Elsevier Science Inc.

Dillenbourg, P., Järvelä, S., & Fischer, F. (2009). The evolution of research on computer-supported collaborative learning. In N. Balacheff, S. Ludvigsen, T. Jong, A. Lazonder, & S. Barnes (Eds.), *Technology-enhanced learning* (pp. 3–19). Dordrecht, Netherlands: Springer. doi:10.1007/978-1-4020-9827-7

Dochy, F., Segers, M., & Buehl, M. M. (1999). The relation between assessment practices and outcomes of studies: The case of research on prior knowledge. *Review of Educational Research, 69*(2), 145–186b. doi:10.3102/00346543069002145

Dubois, D. (2006). Possibility theory and statistical reasoning. *Computational Statistics & Data Analysis, 51*(1), 47–69. doi:10.1016/j.csda.2006.04.015

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*(7), 7–22. doi:10.1007/s10956-008-9119-1

Elinich, K. J. (2011). *Augmented hands-on: An evaluation of the impact of augmented reality technology on informal science learning behavior* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Database. (861926203).

Enyedy, N., Danish, J. A., Delacruz, G., & Kumar, M. (2012). Learning physics through play in an augmented reality environment. *International Journal of Computer-Supported Collaborative Learning, 7*(3), 347–378. doi:10.1007/s11412-012-9150-3

Fotouhi-Ghazvini, F., Earnshaw, R. A., Robison, D., & Excell, P. S. (2009). Designing augmented reality games for mobile learning using an instructional-motivational paradigm. Paper presented at the 2009 International Conference on CyberWorlds (pp. 312–319). Bradford, United Kingdom: IEEE. doi:10.1109/CW.2009.27

Franklin, C., & Garfield, J. B. (2006). The GAISE Project: Developing statistics education guidelines for pre K-12 and college courses. In G. Burrill (Ed.), *Thinking and reasoning with data and chance: 2006 NCTM yearbook* (pp. 375–435). Reston, VA: National Council of Teachers of Mathematics.

Gal, E., & Nachmias, R. (2011). Online learning and performance support in organizational environments using performance support platforms. *Performance Improvement, 50*(8), 25–32. Retrieved from http://login.ezproxy1.lib.asu.edu/login?url=http://search.proquest.com/docview/899207269?accountid=4485

Gal, I., & Garfield, J. B. (Eds.). (1997). Curricular goals and assessment challenges in statistics education. *The assessment challenge in statistics education* (pp. 1–14). Lansdale, PA: IOS Press.

Garfield, J. (1995). How students learn statistics. *International Statistical Review, 63*(1), 25–34. Retrieved from http://www.jstor.org/stable/1403775

Garfield, J. (2002). The challenge of developing statistical reasoning. *Journal of Statistics Education, 10*(3), 58–69. Retrieved from http://www.amstat.org/publications/jse/v10n3/garfield.html

Garfield, J., & Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: Implications for research. *Journal for Research in Mathematics Education, 19*, 44–63. Retrieved from http://www.jstor.org/stable/749110

Garfield, J., & Ben-Zvi, D. (2004). Research on statistical literacy, reasoning, and thinking: Issues, challenges, and implications. In D. Ben-Zvi & J. B. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 397–409). Dordrecht, Netherlands: Kluwer Academic Publishers.

Garfield, J., & Ben-Zvi, D. (2007). How students learn statistics revisited: A Current review of research on teaching and learning statistics. *International Statistical Review, 75*(3), 372–396. doi:10.1111/j.1751-5823.2007.00029.x

Garfield, J., & Ben-Zvi, D. (2008). Developing students' statistical reasoning: Connecting research and teaching practice. New York, NY: Springer.

Garfield, J., delMas, R. C., & Chance, B. (2007). Using students' informal notions of variability to develop an understanding of formal measures of variability. In M. Lovett & P. Shah (Eds.), *Thinking with data* (pp. 117–147). Mahwah, NJ: Lawrence Erlbaum Associates.

Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. New York, NY: Palgrave MacMillan.

Gee, J. P. (2007). Good video games + good learning: Collected essays on video games, learning, and literacy. New York, NY: Peter Lang Publishing.

Gee, J. P. (2011). Reflections on empirical evidence on games and learning. In S. Tobias & J. D. Fletcher (Eds.), *Computer games and instruction* (pp. 223–232). Charlotte, NC: Information Age Publishers.

Giraud, G. (1997). Cooperative learning and statistics instruction. *Journal of Statistics Education, 5*(3), [Online]. Retrieved from http://www.amstat.org/publications/ jse/v5n3/giraud.html

Glenberg, A. M. (2008). Embodiment for education. In P. Calvo (Ed.), *Handbook of cognitive science: An embodied approach* (pp. 355–372). New York, NY: Elsevier Science Limited.

Glenberg, A. M. (2010). Embodiment as a unifying perspective for psychology. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(4), 586–596. doi:10.1002/wcs.55

Glenberg, A., & Glenberg, A. M. Andrzejewski, M. (2007). Learning from data: An introduction to statistical reasoning. Mahwah, NJ: Lawrence Erlbaum Associates.

Gredler, M. E. (2011). Understanding Vygotsky for the classroom: Is it too late? *Educational Psychology Review, 24*(1), 113–131. doi:10.1007/s10648-011-9183-6

Häkkinen, P., & Hämäläinen, R. (2012). Shared and personal learning spaces: Challenges for pedagogical design. *The Internet and Higher Education, 15*(4), 231–236. doi:10.1016/j.iheduc.2011.09.001

Hirsch, L. S., & O'Donnell, A. M. (2001). Representativeness in statistical reasoning: Identifying and assessing misconceptions. *Journal of Statistics Education, 9*(2), 1–21. Retrieved from http://www.amstat.org/publications/jse/v9n2/hirsch.html

Höllerer, T. H., & Feiner, S. K. (2004). Mobile augmented reality. In H. Karimi & A. Hammad (Eds.), *Telegeoinformatics: Location-based computing and services* (pp. 1–39). London: Taylor & Francis Books Ltd. Retrieved from http://www.richiereyes.com/books/Mobile Augmented Reality.pdf

Höllerer, T., Feiner, S., Terauchi, T., Rashid, G., & Hallaway, D. (1999). Exploring MARS: Developing indoor and outdoor user interfaces to a mobile augmented reality system. *Computers & Graphic, 23*(6), 779–785.

Holz, T., Campbell, A. G., O'Hare, G. M. P., Stafford, J. W., Martin, A., & Dragone, M. (2011). MiRA—Mixed Reality Agents. *International Journal of Human-Computer Studies, 69*(4), 251–268. doi:10.1016/j.ijhcs.2010.10.001

Huizenga, J., Admiraal, W., Akkerman, S., & Dam, G. Ten. (2009). Mobile game-based learning in secondary education: Engagement, motivation and learning in a mobile city game. *Journal of Computer Assisted Learning, 25*(4), 332–344. doi:10.1111/j.1365-2729.2009.00316.x

International Telecommunications Union. (2012). Measuring the Information Society 2012. Retrieved from http://www.itu.int/ITU-D/ict/publications/idi/index.html

Jetter, H., Geyer, F., Schwarz, T., & Reiterer, H. (2012). Blended interaction–Toward a framework for the design of interactive spaces. In P. Thomas (Ed) *Proceedings of Workshop on Designing Collaborative Interactive Spaces AVI2012*. Capri Island, Italy: ACM Press. doi: http://doi.acm.org/10.1145/2254556.2254729

Johnson-Glenberg, M. C., Birchfield, D. A., Tolentino, L., & Koziupa, T. (2013). Collaborative embodied learning in mixed reality motion-capture environments: Two science studies. *Journal of Educational Psychology*, Advance online publication. doi: 10.1037/a0034008

Jones, G. A., Langrall, C. W., & Mooney, E. S. (2007). Research in probability: Responding to classroom realities. In F. K. Lester (Ed.), *The Second Handbook of Research on Mathematics* (pp. 909–956). Reston, VA: National Council of Teachers of Mathematics (NCTM).

Kafai, Y. B., Quintero, M., & Feldon, D. (2010). Investigating the "Why" in Whypox: Casual and systematic explorations of a virtual epidemic. *Games and Culture, 5*(1), 116–135. doi:10.1177/1555412009351265

Kahneman, D., Slovic, P., & Tversky, A. (1982). Judgment under uncertainty: Heuristics and biases. New York, NY: Cambridge University Press.

Kamarainen, A. M., Metcalf, S., Grotzer, T., Browne, A., Mazzuca, D., Tutwiler, M. S., & Dede, C. (2013). EcoMOBILE: Integrating augmented reality and probeware

with environmental education field trips. *Computers & Education, 68*, 1–12. Retrieved from http://dx.doi.org/10.1016/j.compedu.2013.02.018

Kaput, J. J. (1992). Technology and mathematics education. In D.A. Grouws (Ed.), *Handbook of teaching and learning mathematics* (pp. 515–556). New York, NY: Palgrave MacMillan.

Katz, N., Cook, T., & Smart, R. (2011). Extending web browsers with a unity 3D-based virtual worlds viewer. *Internet Computing, 15*(5), 15–21. Los Alamitos, CA: IEEE Computer Society. doi:10.1109/MIC.2011.74

Kaufmann, H., & Schmalstieg, D. (2003). Mathematics and geometry education with collaborative augmented reality. *Computers & Graphics, 27*(3), 339–345. doi:10.1016/S0097-8493(03)00028-1

Keeler, C. M., & Steinhorst, R. K. (1995). Using small groups to promote active learning in the introductory statistics course: A report from the field. *Journal of Statistics Education, 3*(2), [Online]. Retrieved from http://www.amstat.org.ezproxy1.lib.asu.edu/PUBLICATIONS/JSE/v3n2/keeler.html

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

Key curriculum. (2013). *Fathom dynamic data* [Software]. Retrieved from http://www.keycurriculum.com/products/fathom

Klopfer, E., & Squire, K. (2008). Environmental detectives—the development of an augmented reality platform for environmental simulations. *Educational Technology Research and Development, 56*(2), 203–228. doi:10.1007/s11423-007-9037-6

Klopfer, E., Squire, K., & Jenkins, H. (2002). Environmental Detectives: PDAs as a window into a virtual simulated world. In M. Milrad & U. Hoppe (Eds.), *Proceedings of the IEEE International Workshop on Wireless and Mobile Technologies in Education* (pp. 95–98). Vaxjo, Sweden: IEEE Computer Society. doi:10.1109/WMTE.2002.1039227

Kreijns, K., Kirschner, P. A., & Jochems, W. (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: A review

of the research. *Computers in Human Behavior, 19*(3), 335–353. doi:10.1016/S0747-5632(02)00057-2

Kreiser, K. (2012, February 07). Intra-operative see-through head mounted displays from the point of view of a designer—a study [Web log post]. Retrieved from http://medicalaugmentedreality.com/2012/02/see-through-head-mounted-displays-from-the-point-of-view-of-a-designer-a-study-by-kilian-kreiser/

Kroeker, K. L. (2010, July). Mainstreaming augmented Reality. *Communications of the ACM, 53*(7), 19–21. doi:10.1145/1785414.1785422

Kurti, A., Milrad, M., & Spikol, D. (2007). Designing innovative learning activities using ubiquitous computing. In *Proceedings of the Seventh IEEE International Conference on Advanced Learning Technologies* (pp. 386–390). Niigata, Japan: IEEE. doi:10.1109/ICALT.2007.118

Lakoff, G., & Johnson, M. (1980). Metaphors we live by. Chicago, IL: University of Chicago Press.

Lane, D. (1999). The Rice virtual lab in statistics. *Behavior Research Methods, Instruments, & Computers, 31*(1). 24-33. doi: 10.3758/BF03207689

Lane, D. M., & Tang, Z. (2000). Effectiveness of simulation training on transfer of statistical concepts. *Journal of Educational Computing Research, 22*(4), 383–396.

Lave, J., & Wenger, E. (1991). Situated learning: Legitimate peripheral participation. New York, NY: Cambridge University Press.

Lin, G.-Y. (2011). Designing a web-based collaborative-learning module for statistical problem solving. *British Journal of Educational Technology, 42*(3), E54–E57. doi:10.1111/j.1467-8535.2011.01172.x

Lipsey, M. W., & Wilson, D. B. (1993) The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist, 48*(12), 1181-1209.

Looi, C.-K., Seow, P., Zhang, B., So, H.-J., Chen, W., & Wong, L.-H. (2010). Leveraging mobile technology for sustainable seamless learning: A research agenda. *British Journal of Educational Technology, 41*(2), 154–169. doi:10.1111/j.1467-8535.2008.00912.x

Lovett, M. (2001). A collaborative convergence on studying reasoning processes: A case study in statistics. In D. Klahr & S. Carver (Eds.), *Cognition and instruction: 25 years of progress* (pp. 347–384). Mahwah, NJ: Lawrence Erlbaum Associates.

Lunsford, M., Holmes-Rowell, G., & Goodson-Espy, T. (2006). Classroom research: Assessment of student understanding of sampling distributions of means and the central limit theorem in post-calculus probability and statistics classes. *Journal of Statistics Education, 14*(3), 1–19.

Magel, R. C. (1998). Using cooperative learning in a large introductory statistics class. *Journal of Statistics Education, 6*(3), [Online].

Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist, 32*(1), 1–19. doi:10.1207/s15326985ep3201

Meletiou-Mavrotheris, M., & Lee, C. (2002). Teaching students the stochastic nature of statistical concepts in an introductory statistics course. *Statistics Education Research Journal, 1*(2), 22–37.

Moore, D. S. (1997). New pedagogy and new content: The case of statistics. *International Statistical Review, 65*(2), 123–165. doi:10.2307/1403333

Morgan, M., & Butler, M. (2009). Considering multi-touch display technology for collaboration in the classroom. In *Proceedings of the World Conference on Educational Multimedia, Hypermedia and Telecommunications 2009* (pp. 674–683). Honolulu, HI: AACE. Retrieved from http://www.editlib.org/p/31573/

Naismith, L., Lonsdale, P., Vavoula, G., & Sharples, M. (2004). *Literature Review in Mobile Technologies and Learning.* Retrieved from http://www2.futurelab.org.uk/resources/documents/lit_reviews/Mobile_Review.pdf

Núñez, R. (2012). On the science of embodied cognition in the 2010s: Research questions, appropriate reductionism, and testable explanations. *Journal of the Learning Sciences, 21*(2), 324–336.

O'Brien, H. L., & Toms, E. G. (2010). The development and evaluation of a survey to measure user engagement in ecommerce environments. *Journal of the American Society for Information Science & Technology, 61*(1), 50–69. doi:10.1002/asi.21229

Park, S. I., Lee, G., & Kim, M. (2009). Do students benefit equally from interactive computer simulations regardless of prior knowledge levels? *Computers & Education*, *52*(3), 649–655. doi:10.1016/j.compedu.2008.11.014

Piaget, J. (1953). The origin of intelligence in the child. London, UK: Routledge & Kegan Paul.

Plass, J. L., Homer, B. D., Milne, C., Jordan, T., Kalyuga, S., Kim, M., & Lee, H. (2011). Design factors for effective science simulations: Representation of information. In R. E. Ferdig (Ed.), *Discoveries in gaming and computer-mediated simulations: New interdisciplinary applications* (pp. 16–35). Hershey, PA: Information Science Reference.

Pratt, D., Davies, N., & Connor, D. (2011). The role of technology in teaching and learning statistics. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics—Challenges for teaching and teacher education* (pp. 97–107). New York, NY: Springer.

Price, S., & Rogers, Y. (2004). Let's get physical: The learning benefits of interacting in digitally augmented physical spaces. *Computers & Education, 43*(1-2), 137–151. doi:10.1016/j.compedu.2003.12.009

Quinn, C. N. (2011). *Designing mLearning: Tapping into the mobile revolution for organizational performance*. San Francisco, CA: Pfeiffer.

Roseth, C. J., Garfield, J. B., & Ben-Zvi, D. (2008). Collaboration in learning and teaching statistics. *Journal of Statistics Education, 16*(1), 1–15.

Rossman, A. J., & Chance, B. L. (2010). *Workshop statistics: Discovery with data JMP companion manual*. Hoboken, NJ: John Wiley & Sons, Inc.

Ryu, H., & Parsons, D. (2009). Designing learning activities with mobile technologies. In H. Ryu & D. Parsons (Eds.), *Innovative mobile learning: Techniques and technologies* (pp. 1–20). Hershey, PA: IGI Global. doi:10.4018/978-1-60566-062-2

Schwabe, G., & Göth, C. (2005). Mobile learning with a mobile game: Design and motivational effects. *Journal of Computer Assisted Learning, 21*(3), 204–216.

Schwabe, G., Göth, C., & Frohberg, D. (2005). Does team size matter in mobile learning? In W. Brookes, E. Lawrence, R. Steele, & E. Chang (Eds.), *Proceedings of the International Conference of Mobile Business* (pp. 227–234). Sydney, Australia: IEEE. doi: 10.1109/ICMB.2005.35

Schwartz, D.L., Sears, D., & Chang, J. (2007). Reconsidering prior knowledge. In M. C. Lovett & P. Shah (Eds.), *Thinking with data* (pp. 319–344). Mahwah, NJ: Lawrence Erlbaum Associates.

Sedlmeier, P. (1999). *Improving statistical reasoning: Theoretical models and practical implications.* New York, NY: Psychology Press.

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

Shaffer, J. P. (1986). Modified sequentially rejective multiple test procedures. *Journal of the American Statistical Association, 81*(395), 826–831.

Shafto, P., & Goodman, N. (2008). Teaching games: Statistical sampling assumptions for learning in pedagogical situations. In B.C. Love, K. McRae, & V. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 1632–1637). Austin, TX: Cognitive Science Society.

Sharples, M. (2000). The design of personal mobile technologies for lifelong learning. *Computers & Education, 34*(3-4), 177–193. doi:10.1016/S0360-1315(99)00044-5

Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In F. K. Lester (Ed.), *The second handbook of research on mathematics teaching and learning* (pp. 957–1010). Reston, VA: National Council of Teachers of Mathematics (NCTM).

Shelton, B. E. (2003). *How augmented reality helps students learn dynamic spatial relationships* (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses Database. (305274787).

Singleton, R. (1989). On teaching sampling: A classroom demonstration of concepts, principles, and techniques. *Teaching Sociology, 17*(3), 351–355.

Snee, R. D. (1993). What's missing in statistical education? *The American Statistician, 47*(2), 149–154.

So, H., Kim, I., & Looi, C.-K. (2008). Seamless mobile learning: Possibilities and challenges arising from the Singapore experience. *Educational Technology International, 9*(2), 97–121.

Squire, K. (2005). Changing the game: What happens when video games enter the classroom. *Innovate: Journal of Online Education, 1*(6), [Online]. Retrieved from http://website.education.wisc.edu/kdsquire/tenure-files/manuscripts/26-innovate.pdf

Squire, K. (2008). Open-ended video games: A model for developing learning for the interactive age. In K. Salen (Ed.), *The ecology of games: Connecting youth, games, and learning* (pp. 167–198). Cambridge, MA: The MIT Press. doi:10.1162/dmal.9780262693646.167

Squire, K. D., & Jan, M. (2007). Mad city mystery: Developing scientific argumentation skills with a place-based augmented reality game on handheld computers. *Journal of Science Education and Technology, 16*(1), 5–29. doi:10.1007/s10956-006-9037-z

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

Starner, T., Mann, S., Rhodes, B., Levine, J., Healey, J., Kirsch, D., … Pentland, A. (1997). Augmented reality through wearable computing. *Presence: Teleoperators and Virtual Environments, 6*(4), 386–398.

Tan, K. T. W., Lewis, E. M., Avis, N. J., & Withers, P. J. (2008). Using augmented reality to promote an understanding of materials science to school children. In *Proceedings of the ACM SIGGRAPH Asia Educators Programme*. Singapore, Asia: ACM. doi:10.1145/1507713.1507716

Varela, F. J., Thompson, E. T., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge, MA: The MIT Press.

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

Wagner, Daniel, Pintaric, T., Ledermann, F., & Schmalstieg, D. (2005). Towards massively multi-user augmented reality on handheld devices. Paper presented at

3rd International Conference on Pervasive Computing, München, DE, 77–95. Retrieved from https://www.ims.tuwien.ac.at/publications/tuw-139804

Wagner, D., & Schmalstieg, D. (2003). First steps towards handheld augmented reality. In S. Feiner & D Mizell (Eds.), *Proceedings of the Seventh IEEE International Symposium on Wearable Computers* (pp. 127–135). White Plains, NY: IEEE. Retrieved from http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1241402&isnumber=27832

Watkins, J., & Mazur, E. (2013). Retaining Students in Science, Technology, Engineering, and Mathematics (STEM) Majors. *Journal of College Science Teaching*, *42*(5), 36-41. Retrieved from http://www.nsta.org/publications/article.aspx?id=najfhg4vZwI=

Webster, J., & Ho, H. (1997). Audience engagement in multimedia presentations. *ACM SIGMIS Database, 28*(2), 63–77. doi:10.1145/264701.264706

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9(*4), 625–636. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12613670

Xu, Y., Barba, E., Radu, I., Gandy, M., Schrank, B., Macintyre, B., & Tseng, T. (2011). Pre-patterns for designing embodied interactions in handheld augmented reality games. In *Proceedings of the* 2011 IEEE International Symposium on Mixed and Augmented Reality--Arts, Media, and Humanities (pp. 19–28). Switzerland, Austria: IEEE.

Yuen, S. C., Yaoyuneyong, G., & Johnson, E. (2011). Augmented reality: An overview and five directions for AR in education. *Journal of Educational Technology Development and Exchange, 4*(1), 119–140.

APPENDIX A

PRESURVEY

Pre-survey: Your responses for this 5 minute survey are anonymous. Please answer the questions as truthfully as you can.

1. Enter your Study ID number.

2. What is your gender? (Male or Female)

3. How old are you?

4. What year are you in school? (Freshman, Sophomore, Junior, or Senior)

5. Have you ever taken a statistics course in high school or college? (Yes or No)

6. How many statistics course have your taken since high school? (0, 1-2, 3-4, or 5 or more)

7. How would you rate your ability in statistics? (Excellent, Good, Fair, or Poor)

8. How would you rate your attitude towards statistics? (Excellent, Good, Fair, or Poor)

9. How would you rate your ability in calculating proportions? (Excellent, Good, Fair, or Poor)

10. How would you rate your ability in calculating sampling variability? (Excellent, Good, Fair, or Poor

11. How would you rate your ability to use mobile devices (i.e. smartphone, tablets, laptops)? (Excellent, Good, Fair, or Poor)

12. How would you rate your ability to use iPads? (Excellent, Good, Fair, or Poor)

APPENDIX B

PRETEST AND POSTTEST

Pre- and Post-test: *Welcome to the MARVEL research project! This is a survey about statistics and the environment. Your answers will help us with the design of a new way to learn about statistics and the environment. This survey is not part of your grade, and your answers will be used confidentially. You will be asked these same questions again at the end of the MARVEL project. The comparison between the pre- and post-surveys will help us understand whether MARVEL worked well for you or not. Thank you for your time and your help on this very important project! -- The MARVEL Research Team*

Your responses for this 10 minute survey are anonymous. Please answer the questions as best you can.

Obj 1: Build and describe sample distributions
1. A _____ is a set of numerical data. It usually includes results from a survey or experiment.
    a. **Sample**
    b. Random Sampling
    c. Probability
    d. Population

2. _____ is defined by assigning equal probabilities to each possible sample.
    a. Sample
    b. **Random Sampling**
    c. Probability
    d. Population

3. What is a sampling distribution?
    a. Formally, a sampling distribution of a statistic is the probability distribution of the statistic computed from all possible random samples of the same size from the same population.
    b. A sampling distribution is a probability distribution that infers the likihood of some measure. They show the probability of a statistic.
    c. It is a distribution of statistics obtained by selecting all the possible samples of a specific size from a population.
    d. **All of the above.**

Obj 2: Calculate a probability statistic of a sample
4. _____ is a number expressing the likelihood that a specific event will occur, expressed as the ratio of the number of actual occurrences to the number of possible occurrences.
    a. Sample
    b. Random Sampling
    c. **Probability**
    d. Population
5. A standard deck of playing cards has 52 cards divided equally into four suits. Two suits are red (hearts and diamonds), and two are black (clubs and spades). Within each suite are cards labeled 1 (ace) to 10, jack, queen, and king. Given random sampling of a single card from a deck, what is the approximate chance of drawing the ace of diamonds on the first draw?
    a. **2%**
    b. 10%
    c. 25%
    d. 50%
6. If you draw a single card once from a standard deck of 52 playing cards, what is the probability that it will be the ace of diamonds?
    a. **2%**
    b. 10%
    c. 25%
    d. 50%

7. At the beginning of the baseball season in a particular year, the New York Yankees have a 33% chance to win the American League pennant. What is the probability that the Yankees will win the pennant? (Answer: 33% or .3333)

Obj 3: Recognize the variability between samples
8. _____ is the extent to which the measurements in a sampling distribution differ from one another.

      a.   Sampling distribution
      b.   Variability
      c.   Statistic
      d.   Average

Obj 4: Describe the effect of sample size on how well a sample resembles a population

9. A _____ is the set of all the individuals of interest in a particular study.
      a.   Sample
      b.   Random Sampling
      c.   Probability
      d.   Population

10. What happens to a PROPORTION of sample if they the number of fish in each sample (sample size) increases?
      a.   The proportion of the sample gets closer to the proportion of the population
      b.   The proportion of the sample gets further from the proportion of the population

11. If you take 10 separate samples of fish of the same size from a pond, is the PROBABILITY getting closer or further from the PROBABILITY found in the population?
      a.   Closer
      b.   Further
      c.   Not sure

12. **True** or False: The larger the sample size, the better the sample resembles a population.

Obj 5: Describe the differences between sample and population distributions

13. Which of the following statements is an example of sample.
      a.   Number smartphone users in the entire world.
      b.   The number of days spent in intensive care for all people who have undergone heart transplant surgery.
      c.   The number of words recalled from a list of 50 words by 25 first-year college students who volunteer to take part in an experiment.

14. Which of the following statements is an example of a population.
      a.   The number of words recalled from a list of 50 words by 25 first-year college students who volunteer to take part in an experiment.
      b.   The number of days spent in intensive care for all people who have undergone heart transplant surgery.
      c.   The number of errors made by rats learning a maze.

15. Given a sample of 100 fish from Pond A, if the probability of Pond A is .368, what percentage of all the fish do you predict are affected?
      a.   2%
      b.   11%
      c.   37%
      d.   63%

16. As you increase your sample size, how will the proportion of your sample that is affected by the pollution relate to the proportion of the whole population that is affected?
      a.   I expect the sample proportion to be the LESS like the POPULATION proportion.
      b.   I expect the sample proportion to be the MORE like the POPULATION proportion.
      I expect the sample proportion to be UNRELATED to the POPULATION proportion

APPENDIX C

POSTTEST ONLY

Post-test Only: *For the following questions, read the question carefully and answer to the best of your ability.*

1. Based on the results you found during the experimental session, what type of oil do you suspect is polluting the water?

   a. Class A Oils:
      - In water, class A oils disperse readily but affect aquatic life in the upper water column.
      - Class A oils include high-quality refined fuels such as gasoline, kerosene, and jet fuel.
      - Are light and fluid, spread quickly when spilled and have a strong odor.
      - Are the most toxic but least persistent of all oils.
   b. Class B Oils:
      - Class B oils include most types of crude oils. Such oils are prone to forming lumps of oil or sludge.
      - Class B oils do not easily dilute and disperse, making it especially detrimental to wildlife.
      - Class B oils are sticky and adhere strongly to objects it comes into contact with.
      - Are heavy and have the tendency to sink in water over time.
   c. Non-Petroleum Oils:
      - Are synthetic oils, derived from plant or animal fats, for purposes such as cooking or lubricants.
      - Non-petroleum oils are slow to break down, causing long-lasting damage to an affected area.
      - Non-petroleum oils coat wildlife and can cause death due to suffocation or dehydration.
      - Examples of non-petroleum oil products include cooking fats and vegetable oils.

2. What do you think the source of the pollutant is?

   a. A fuel spillage from a nearby oil refinery located by the shore.
   b. An oil tanker leak out in the open water that hasn't been contained yet.
   c. The public landfill run off into a neighboring tributary that connects with this bay.

3. Given the suspected source and type of oil spilled, which of the following cleanup methods would you recommend?

   a. Flushing: can be effective in areas where the water is shallow and near shore.
   b. Dredging: cleaning deep below the surface of the impacted water.
   c. Bioremediation: using microorganisms or biological agents to break down and disolve oils.

APPENDIX D

PERCEPTION OF COLLABORATION SCALE

Collaboration: *Please answer the following questions on a scale of 1-10, (1 = strongly disagree, 10 = strongly agree).*

Communication	My partner and I participated equally to complete the learning task.
	It was easy to work together during the learning task.
	I communicated with my partner a lot during the learning task.
Sharing of the Device	My partner and I took turns using the iPad to complete the learning.
	I used the iPad about the same amount as my partner did.
Participation	My partner and I participated equally to complete the learning task.
	My partner and I had the same level of knowledge.
Cooperation	My partner helped me during the learning activity.
	I helped my partner during the learning activity.

APPENDIX E

PERCEPTION OF ENGAGEMENT SCALE

Engagement: *Please answer the following questions on a scale of 1-7, (1 = strongly disagree, 10 = strongly agree).*

Focused Attention

I forgot about my immediate surroundings while using the app.
I was so involved in the learning task that I ignored everything around me.
I lost myself in this learning experience.
I was so involved in my learning task that I lost track of time.
I blocked out things around me when I was using the app.
When I was using the app, I lost track of this world around me.
The time I spent using the app just slipped away.
I was absorbed in the learning task.
During this learning experience I let myself go.

Perceived Usability

I felt frustrated while using this app.
I found this app confusing to use.
I felt annoyed while using this app.
I felt discouraged while using this app.
Using this app was mentally taxing.
This learning experience was demanding.
I felt in control of my learning experience.
I could not do some of the things I needed to do on this app.

Aesthetics

This app is attractive.
This app was aesthetically appealing.
I liked the graphics and images used on this app.
This app appealed to my visual senses.
The screen layout of this app was visually pleasing.

Endurability

Learning on this app was worthwhile.
I consider my learning experience a success.
This learning experience did not work out the way I had planned.
My learning experience was rewarding.
I would recommend using this app to my friends and family.

Novelty

I continued to use this app out of curiosity.
The content of the app incited my curiosity.
I felt interested in my learning task.

Involvement

I was really drawn into my learning task.
I felt involved in this learning task.
This learning experience was fun.